5-2004

Pharmacological and Behavioral Treatments for Migraine Headaches: A Meta-Analytic Review

Kristi Lowe Stewart
Utah State University

PHARMACOLOGICAL AND BEHAVIORAL TREATMENTS FOR MIGRAINE HEADACHES: A META-ANALYTIC REVIEW

by

Kristi Lowe Stewart

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

2004

ABSTRACT

Pharmacological and Behavioral Treatments for

Migraine Headaches: A Meta-Analytic Review

by

Kristi Lowe Stewart, Doctor of Philosophy

Utah State University, 2004

Major Professor: Dr. Kevin Masters
Department: Psychology

Migraine headache is a painful and often serious problem in the United States. There are many prophylactic pharmacological and nonpharmacological treatments available for migraine headaches. However, choosing between them can be difficult for individuals and treatment providers alike. The primary literature regarding the effectiveness of pharmacological and nonpharmacological treatments is quite dense. More than 191 primary studies were identified as providing information about the effectiveness of one or more treatments for migraine headaches. Of these, 82 articles were retained for meta-analyses on six prophylactic treatments for migraine: propranolol, flunarizine, divalproex sodium, thermal biofeedback, relaxation treatments, and combined treatments. These results suggest that all of the treatments examined have very similar treatment effect sizes. Effect sizes for the reduction of migraine were found to be between .60 and .75 for

all treatments. Results suggest that selection of treatment for migraine must be based on more than treatment effectiveness rates alone.

(200 pages)

DEDICATION

This document is dedicated to my friends and family who have supported me throughout my life. The time and energy that I have devoted to my graduate studies has often been at the expense of those I love. This project could never have been completed without the support of Dr. Kevin Masters, who never gave up on me or this project. Of particular importance to me has been the unwavering love and support of David Shearer and the guidance of my family through the many years it took to complete this project. I wish to give special thanks to Cynthia Cole, Scott Hill, Richard Farley, Raylene Glover, Lara and Rick LaCaille, and Scott Stewart for their encouragement and support while writing my dissertation.

Kristi Lowe Stewart

CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

PROBLEM STATEMENT

Recurrent headaches are a problem for millions of men and women (Goldstien & Chen, 1982). One type of headache, the migraine, is a disabling syndrome resulting in significant life impairment (Unruh, 1996). There are two categories of treatments for migraine headaches. The first, pharmacological treatments, encompasses several classes of agents that are used in the treatment of headache pain including beta-adrenergic blockers, calcium channel blockers, and anticonvulsants (Solomon, 1995). The second category of treatments for migraine headaches are nonpharmacological. These include thermal and electromyograph (EMG) biofeedback, relaxation training, and cognitive techniques (Capobianco, Cheshire, & Campell, 1996).

The primary literature regarding the effectiveness of pharmacological and nonpharmacological treatments is quite dense. More than 191 primary studies were identified as providing information about the effectiveness of one or more treatments for migraine headaches. A majority of these primary studies (approximately 65%) compared their active treatment to a placebo or control group. Of those that did compare to other active treatments only three compared pharmacological and nonpharmacological treatments in the same study. While control group research comparisons give valuable information regarding whether the treatments are better than no treatment, this type of research design does little to inform about comparative effectiveness. As a result, there are at least 16 different categories of treatments with some empirical evidence supporting their use, but little data indicating which are the "better" treatments. Obviously, sorting through

this amount of information and selecting the most effective treatment can be overwhelming.

Attempts have been made to synthesize the effectiveness literature for migraine headaches through reviews. However, most of the existing reviews narrowly focus on either pharmacological or nonpharmacological treatments. There are two existing reviews that evaluated both pharmacological and nonpharmacological treatments for migraine headaches in adult populations (Holroyd & Penzien, 1990; Duke University and the Center for Clinical Health Policy Research, 1999a, 1999b). The first (Holroyd & Penzien, 1990) examined 73 primary studies, however, it examined only one pharmacological treatment (Propranonol) and two nonpharmacological treatments (relaxation and relaxation/biofeedback). Statistical methods (percent improvement) used in this meta-analysis are unfocused estimations of effect and are not standardized. In addition, since the Holroyd review (which collected data through 1989), there have been several new studies published that examine the effectiveness of treatments for migraine headaches.

The second review consisted of two sections of review developed by the Agency for Health Care Policy and Research (Duke University, Center for Clinical Health and Policy Research, 1999a, 1999b). This review was published after the proposal date of the current project and it mirrors many of the current authors intentions. The first part of the Duke University (1999a) review synthesized the data for behavioral and physical treatments for migraine. This methodologically sophisticated review was excellent in most aspects. The second part of the review (1999b) synthesized the data for prophylactic pharmacological agents for migraine. The review only included randomized controlled

trials of pharmacological, psychological, behavioral and physical (e.g., acupuncture) treatments. However, this meta-analysis did not compare the findings between pharmacological and nonpharmacological groups. Of interest was the authors' choice in this review to include the percent improvement measure in the nonpharmacological data and not in the pharmacological data. This was curious given their emphasis on consistency of evaluation methods across domains. In addition, the results have not been published (likely due to length and complexity) outside of the National Technical Information Service (NTIS). Therefore, the area of migraine headaches is without an easily accessible comprehensive review of the primary research literature.

The lack of synthesis of the literature between two supposedly empirically supported categories of treatment made it difficult to determine which, if any, of the treatments were most effective. It is important to understand how pharmacological and nonpharmacological treatments compare to one another so that practitioners and consumers can make informed choices. Therefore, a new meta-analytic review needed to be conducted that integrated the data from both pharmacological and nonpharmacological research. A single integrative review of the effectiveness findings for migraine will serve as a valuable resource for information regarding recurrent headaches and will fill a void that exists in the research literature. This project endeavored to fill this void by conducting meta-analyses on six prophylactic treatments for migraine and comparing them to one another.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

This review of the literature begins with a brief overview of migraine headaches, other related recurrent headaches syndromes, and their treatments. Recurrent headaches are a problem for millions of Americans and many treatment options are available. However, this review of migraine headaches and the two treatment categories available (pharmacological and nonpharmacological) will illustrate that the sheer number of available treatment choices can create decision-making difficulties for professionals and consumers alike. It is suggested that what is needed in the area of migraine headache treatment research is a synthesis of the data and a comparison of effectiveness rates across categories. One option to meet this need, a meta-analysis, is presented. Previous reviews on the topic are then explored to examine where previous meta-analyses have succeeded and failed. In addition, previous reviews are examined to develop guidelines for a proposed meta-analytic review.

Recurrent Headache Syndromes

While most people experience mild headaches from time to time, millions of people each year are affected by disabling headaches (Holroyd & French, 1995). New cases of disabling recurrent headaches are estimated to occur at nearly two million per year; making recurrent headaches a serious health problem in the United States (Goldstein &

Chen, 1982). Recurrent headaches are not a homogenous group of symptoms. Rather, a recent classification system identified 12 diagnostic categories with 145 subcategories of headaches (Headache Classification Committee of the International Headache Society [IHS], 1988). The 12 major diagnostic categories can be broken into three basic types of headaches: migraine, tension, and mixed type. Each type of headache is associated with specific symptom profiles. These profiles help make headache syndromes distinct from one another. This review is primarily interested in exploring migraine headache syndrome. Therefore, this section presents the symptom profile, the prevalence, and the etiology of migraine headache syndrome. It is necessary, however, to establish a general understanding of the full range of other recurrent headache syndromes (tension and mixed) to evaluate how migraine headaches relate to these other syndromes. Thus, tension headache and mixed-type headache will also be briefly described in this section. Then in the following section, treatment modalities for migraine headache syndrome will be explored.

Migraine Headaches

Migraine headaches are characterized by severe periodic pulsating headaches which can range in duration from a few minutes to three days. Migraines have unilateral locations (occur on one side of the head) and are often associated with severe pain, light avoidance, sound avoidance, and vomiting (Taylor, 1991). Migraines have been noted to occur in distinct phases (Blau, 1980). The phases that have typically been identified are the prodrome, aura, headache, resolution, and postdrome. The prodrome phase can occur several hours or days before the onset of the headache phase and can serve as a warning of

an oncoming migraine (Capobianco et al., 1996). Symptoms of the prodrome phase include osmorphobia, photophobia, drowsiness, euphoria, irritability, and food cravings. The aura phase occurs immediately prior to the migraine attack. The symptoms associated with the aura typically last less than an hour and consist of visual disturbances, numbness in limbs, and/or language disorders (Capobianco et al.). The headache phase is typically characterized by unilateral pulsating head pain that lasts for 4 to 72 hours (Holyroyd & French, 1995). Pain diminishes during the resolution phase and during the postdrome phase pain is eliminated and all symptoms vanish. The term migraine refers to a headache that includes any or all of these phases. It is customary to distinguish between "classic migraines," which include a prodrome or aura phase, and "common migraines," which do not.

Migraine suffers often experience their first episode in childhood or adolescence, although onset in the 20s or 3-s is not uncommon. The later the onset of the headache the more likely it is to be associated with organic cause (e.g., cerebrovascular disease or tumor; Dalessio, 1994). In a review of the literature on migraine prevalence Stewart, Schecter, and Rasmussen (1994) reported that migraine occurs most commonly in individuals between the ages of 25 and 50. The authors also reported that lifetime prevalence rates have been reported in the literature to range from 8-42% of the population. Estimates of prevalence vary widely by how migraine headache is defined and diagnosed. According to Stewart et al., studies using the clearly defined IHS criteria for the diagnosis of migraine typically estimate the lifetime prevalence of migraine to be between 24-30%. Migraines are also associated with severe life impairment. A recent

review identified that, on average, researchers have found that 50% of individuals with migraine headaches miss more than one day of work a year due to headache pain, 50% discontinue normal activities, and 30% cancel family and social activities (Stewart et al.). Although the economic impact of a migraine headache is difficult to assess, best estimates indicate the indirect cost of migraines in the United States to be between $1.4 billion-17.2 billion a year (de Lissovoy & Lazarus, 1994). While many of the prevalence and impact estimates are complicated by migraine classification problems, the above estimates suggest that migraine headaches cause serious disruption in the lives of millions of people each year.

There are a variety of current theories that attempt to explain the pathophysiology of migraine. Each theory is extensive and highly debated in the research literature. A full description of each theory is beyond the scope of this paper. However, several of the most influential theories will be briefly discussed. Historically, research on the pathophysiology of migraine headaches has largely been driven by the theories generated by Wolff in the 1940s. Wolff purposed that migraine headaches were caused by increased pressure on major cranial arteries. Increased pressure was hypothesized to be the result of prolonged dilation of veins and arteries (Taylor, 1991). Dilation was suggested to be triggered by an external event. Common events associated with the onset of a migraine have been noted to be psychological stress, orgasm, flickering lights, certain foods (e.g., chocolate and red wine), and sound.

New theories about pathophysiology of migraines have moved beyond Wolff's original conception and have begun to explore the biological mechanisms involved in

migraine and suggest that the pathophysiology of migraine is much more complicated than previously thought. For example, central pain mechanisms have been implicated in the production and maintenance of migraines (Raskin, 1988). Specifically, it has been hypothesized that abnormalities may occur in the brain stem pain modulation system. These abnormalities may cause a rapid depolarization of cerebral neurons and may occur during or just prior to the onset of the migraine headache. Others have suggested that individuals who experience migraine headaches have an inherited susceptibility to headache (Lance, 1993). This susceptibility is thought to be influenced by an interaction between exogenous and endogenous factors. Once influenced (by either an internal event or external event) the balance between inhibitory and excitatory neurons may be altered, which is hypothesized to impact the serotonergic system, vascular pressure, and cranial circulation; thereby, producing intense head pain. While most of the new research on the mechanisms involved in migraine headache has focused in the biological arena, no researchers have ruled out or adequately explained the influence of exogenous factors. New theories for migraine headaches are working to identify the specific mechanisms involved in the production and maintenance of the migraine.

In sum, the course of a migraine headache is idiosyncratic and can often be difficult to diagnose. Symptom profiles can include intense unilateral head pain, vomiting, and photophobia. Many individuals with migraine headaches experience an early onset of pain (childhood or adolescence), but, a later onset in not uncommon. Five distinct phases have been identified in the course of a migraine (prodrome, aura, headache, resolution, and postdrome). Migraine course is not consistent from individual to individual, thus, any or

all of the phases may be present. The pathophysiology of the migraine headache is considered to be highly complex and multifaceted. Current theories consider both external (triggers) and internal events (biological mechanisms) to be important in the etiology of a migraine headache (Capobianco et al., 1996).

Tension Headaches

In contrast to migraine headache, tension headaches are characterized by a bilateral head pain that is nonpulsing and can last up to 7 days. Head pain in tension headaches is usually described as dull, nagging, persistent, tight and constricting (Dalessio, 1994). However, the most severe pain in tension headache may occur in the neck and shoulders. Tension headaches are not accompanied by nausea, photophobia or phonophobia (Holroyd & French, 1995). It has been commonly thought that tension headaches are the result of long-term contraction of cranial and neck muscles (Featherstone, 1985). In addition to muscle contraction, some of the same central pain mechanisms that are being investigated in migraine headaches have also been implicated in the pathophysiology of tension headaches (Raskin, 1988). The prevalence of tension headaches is even more difficult to estimate than prevalence of migraine. It has been suggested that the difficulty in estimating tension headache syndromes lies in the high levels of self-treatment that occurs (over-the-counter medication) and the high degree of overlap that is seen between tension headache pain and other clinical syndromes (migraines, back pain, surgical pain; Featherstone). Despite these complicating factors, it has been suggested that tension

headaches (as defined by the IHS criteria) occur in 88% of women and 69% of men (Rasmussen, Jensen, & Olesen, 1991).

Mixed Headaches

Mixed headache syndromes are those that consist of discrete periods of both tension and migraine headaches. Individuals with mixed headaches may experience both migraine and tension headaches within a 1-month period. Reliable estimates of prevalence and impact of mixed headache syndromes are not available due to problems with diagnosis and research methodology.

Thus, recurrent headaches are a problem for millions of people each year. The migraine is a particularly problematic type of recurrent headache. Migraines result in severe head pain, lost work days, and health care costs that exceed a billion dollars a year nationally. Migraine headache syndrome is differentiated from tension headache syndrome by the type of head pain, the associated symptoms, and the postulated pathophysiology of the headache. When migraine headaches and tension headaches both occur in the same person, the headache syndrome is classified as mixed type. There is a lack of understanding of the etiologic pathways involved in the onset and maintenance of migraine headaches and, currently, there is no permanent cure for migraine headaches. However, there has been a variety of techniques and pharmacological agents developed that are designed to treat migraine headache syndrome. The next section examines these treatments.

Types of Treatment Currently Being Used

The ambiguity surrounding the pathophysiology of migraine headache syndrome has contributed to the proliferation of treatment techniques in both medical and psychological fields. As one might suspect, medical developments have been largely pharmacological in nature; psychological treatments have largely focused on relaxation, biofeedback, and cognitive interventions. Pharmacological treatments are numerous and varied, thus three major categories and the most popular treatment in each category will be briefly reviewed. Then, nonpharmacological treatments will be summarized. The purpose of this section is to identify major treatment categories. The results of the previous outcome literature will be reviewed later.

Pharmacological Treatments

Pharmacological treatments for migraine headaches have evolved significantly over the past two decades. Many narrative reviews published in the past few years attempted to identify possible medication choices for individuals suffering from migraine headaches (e.g., Capobianco et al., 1996; Solomon, 1995). A brief review of available pharmacological agents is provided here (based on Capobianco et al.; Solomon) in order to identify the pharmacological modalities used in the treatment of migraine headaches.

Pharmacological treatments have generally been classified as prophylactic and symptomatic treatments (Capobianco et al., 1996). Symptomatic treatments treat the pain once the symptoms have begun (Solomon, 1995). Symptomatic treatments are commonly used with individuals who have infrequent migraine attacks. Prophylactic treatments are

the primary concern of this paper, prophylactic therapy is used with patients who have regular migraine headaches and works to prevent the onset of the migraine. Capobianco et al. report that a prophylactic treatment should only be considered when any of the following criteria are met: (a) patients are taking excessive amounts or medication, (b) they have regular headaches, (c) they have attacks that are severe and last more than 48 hours, (d) the patient is unable to cope with pain and life effects, and/or, (e) attacks occur after a prolonged aura. When a client meets one or more of these conditions, several classes of medications can be considered.

There are several major pharmaceutical classes of prophylactic agents. The three major classes considered in this project are the beta-adrenergic blockers, calcium channel blockers, and anticonvulsants. These three categories are the most commonly suggested for use by researchers and physicians (Capobianco et al., 1996) and have the most published work on them. Within each class a single agent was chosen for analysis in this paper. The agent for analysis was chosen by identifying the one in each class that had the most published research (for more details see Methods section). Through this fashion, Propranonol (a beta-adrenergic blocker), Flunarizine (calcium channel blocker), and Divalproex Sodium (anticonvulsant) were chosen as the primary targets. Each of the agents have brand names, generic names and chemical names (see Appendix A for list of agent names). For the purposes of this paper the generic names will be used to avoid confusion. A brief review of the properties of each agent follows.

Propranonol. Propranonol, also known as Propranolol Hydrochloride, has been used for migraine prophylaxis for over 30 years. Early reports indicate that practitioners

observed that patients receiving Propranonol for cardiovascular disorders also displayed a decrease in co-existing migraines (e.g., Rabkin, Stables, Levin, & Suzman, 1966). These observations led to a proliferation of research studies on the use of Propranonol for migraine prophylaxis.

Propranonol is known to have a beta-adrenergic blocking effect that is associated with reduced heart rate and blood pressure (Cortelli et al., 1985). It has been suggested that this beta-adrenergic blocking interferes with peripheral vasodilator receptors, thereby restricting the vasodilation phase of a headache (Cortelli et al.). It has also been suggested that Propranonol reduces brain wave activity, thus "quieting" the brain response (Schellenberg, Milch, Schwarz, & Schoberg, 1994). However, not all beta-adrenergic blockers are effective as migraine prophylaxes. Therefore, the mechanisms that account for Propranonol's effects are still largely unknown (Gerber et al., 1995).

Caroll, Reidy, Savundra, Cleave, and McAish (1990) reported that Propranonol has a short half-life ranging from 2-4 hours in its conventional preparation. They go on to state that a long acting formula is available which has a half life of 10-20 hours. Recommended dosages for migraine prophylaxis range from 40-240 mg. Side effects include tiredness, reduced heart rate, reduction of blood pressure, and dizziness (Caroll et al.).

Flunarizine. Flunarizine as a migraine prophylactic has a unique history. Most other migraine prophylactic drugs have been extended to migraine patients after an effect was noted while treating another condition. Flunarizine does not follow this historical

course. Flunarizine is one of the few agents purposed through the scientific method to be an effective prophylactic agent for migraine.

Flunarizine was suggested to be a migraine prophylactic by a group of pharmacists (Amery, 1983) after studying the biopathophysiology of migraine. The Amery group hypothesized that a period of brain hypoxia was the critical event in migraine genesis. Cerebral hypoxia has been associated with the flooding of calcium ions in cerebral arteries (Amery). It has been suggested that cerebral arteries are unique in that they require external calcium ions to contract (as opposed to other vasculature that use intracellular calcium stores to contract). According to the Amery theory, only a drug that decreased the potential for brain hypoxia would be an effective prophylactic agent. Thus, they proposed Flunarizine (a unique calcium channel agonist) would be the most ideal agent available. Flunarizine is a difluorated piperazine derivative that inhibits the influx of calcium ion into vascular smooth muscle cells (Louis, 1981). Thus, the use of Flunarizine may inhibit the influx of the intercellular calcium ion, thereby reducing the likelihood that the cerebral arteries can begin the contraction series seen in migraine (Steardo et al., 1986).

Diamond and Schenbaum (1983) suggested that Flunarizine has another property that contributes to its success as a migraine prophylactic agent. Research indicates that Flunarizine works to block H1 histamine receptors while simultaneously inhibiting vasoconstriction and serotonin activity (Diamond & Schenbaum). Histamine and serotonin have been shown to be present in the brain in high amounts during migraine activity (Steardo et al., 1986).

Flunarizine has a long half-life (duration of action is approximately 24 hours), which results in a delay in treatment effect (Mendopoulous, Manafi, Logothetis, & Bostantzopoulou, 1985). Initial effects are noted between 30 and 90 days (Louis, 1981). Recommended doses for migraine prophylaxis range between 5-10mg. High doses of Flunarizine (more than 30 mg) over long periods of time have been associated with reversible Parkinson-like side effects (Schmidt & Oestreich, 1991). More common side effects include: daytime sedation (Steardo et al., 1986), weight gain (Amery, 1982), depression (Schmidt & Oestreich), dry mouth and skin rash (Martinez-Lage, 1988).

Flunarizine is not currently approved by the FDA for use with migraine patients in the United States. The reason for this is unknown. The researcher contacted the FDA regarding Flunarizine and they responded by stating that it is illegal for them to disclose information about an agent under law 21 CFR 314.430. They stated that all information belongs to the manufacturer of the agent. Thus, I contacted Janssen-Cilag on four occasions to try to get information regarding Flunarizine's drug status in the United States. No response was ever received. Janssen-Cilag is a subsidiary of Johnson and Johnson. Johnson and Johnson moved many of its production plants to Canada and Belgium in the early 1990s. This would have been around the same time the Flunarizine would have become a candidate for FDA review. Even without FDA approval, Flunarizine is one of the most commonly used and researched prophylactic agent in countries outside of the United States.

Divalproex Sodium. This treatment is composed of one part sodium valproate and one part valproic acid (Mathew, 2001). Divalproex Sodium dissociates in the

gastrointestinal tract to valproate. Valproate is hypothesized to increase the levels of gamma-aminobutyric acid (GABA) in the brain (Mathew). It is thought that the increase in GABA levels in the brain serve to reduce the "hyper-excitability" of the brain that is thought to be associated with migraine (Shelton & Connelly, 1996). That is, it is believed that the brains of individuals with migraines are susceptible to a variety of stressors (both physical changes and outside stressors). This is believed to cause the brain to be hyperexcitable and cause a vasoconstriction reaction that will start a migraine.

Divalproex Sodium is also thought to have additional properties that aide in migraine prophylaxis. It is suggested that Divalproex Sodium may inhibit the release of prolactin (Jensen, Brink, & Olesen, 1994). Prolactin is thought to be released in high levels during an ischemia (Shelton & Connelly, 1996). Thus, it is thought that Divalproex Sodium may offer some protection against migraine during the aural phase.

Divalproex Sodium has been used for epilepsy, infantile spasm, photosensitive epilepsy, and migraine (Silberstein & Wilmore, 1966). The FDA has approved Divalproex sodium for use as a migraine prophylaxis. The plasma half life of Divalproex Sodium is 8-17 hours (Silberstien & Wilmore). Recommended doses for migraine prophylaxis are 400-2500 mg a day and migraine reduction is likely to be seen in 30 days (Landy & McGinnis, 1999). Most of the agent is metabolized in the liver, which can result in liver damage if appropriate levels are exceeded (Shelton & Connelly, 1996). Documented side effects can include tremor, weight gain, alopecia (Thomas, 1989), gastrointestinal disorders, and gastroesophageal reflux disease (GERD; Landy & McGinnis).

Nonpharmacological Treatments

A variety of nonpharmacological treatments have been developed to treat migraine headache syndrome. These treatments are based on the theory that controllable responses like stress, tension, heart rate, blood pressure, and negative cognitions contribute to the development, maintenance, and interpretation of the recurrent migraine. The most commonly researched nonpharmacological treatments are briefly reviewed here. It is noted that each treatment described is rooted in a theoretical orientation, has extensive treatment procedures, and has many treatment variations. However, a full discussion of all of these issues is beyond the scope of this paper.

Biofeedback treatments. In a recent nonquantitative review of supported psychological treatments for migraines, Holroyd and French (1995) report that thermal biofeedback plus relaxation training and EMG biofeedback plus relaxation training are two of the most widely used nonpharmacological treatments for migraine headaches. Biofeedback is a method of achieving control over bodily processes that are usually outside of the client's awareness. Biofeedback targets a function (hand warming in thermal feedback and muscle tension in EMG biofeedback) and provides continuous feedback (typically visual and/or auditory) regarding the target function to the client (Taylor, 1991). With continuous feedback, typically provided via a computer, and through trial and error, clients learn to gain control over their target functions. After successful training clients are theoretically able to control their target function without the use of the mechanical feedback (Blanchard, 1994).

Relaxation treatments. Relaxation training is also used to treat clients who suffer from migraine headaches. Relaxation training guides the client to move toward a less physiologically aroused state (Taylor, 1991). Procedures for relaxation training may use any combination of deep breathing techniques, muscle relaxation, and imagery (Malone & Strube, 1988). Treatments typically involve five to eight sessions and include some form of daily home practices. Clients may be given workbooks or audiotapes to aide in their practice at home. In some cases, relaxation techniques have been used as the sole treatment or in combination with biofeedback.

Cognitive treatments. Another commonly reported treatment for migraine headache sufferers is cognitive or cognitive-behavioral therapy. Cognitive treatments are aimed at helping clients alter their perception of pain. In cognitive treatments, clients meet with a therapist and learn to monitor the thoughts, feelings, and behaviors that enhance their sensation of pain (Holroyd & French, 1995). Clients then learn to alter their behavior and thinking to patterns that are associated with less pain. Common protocols for migraine prophylaxis that use cognitive training use components of rational emotive therapy and coping skills training.

Summary

In summation, none of the above described treatments are considered to be a "cure" for migraine headaches. Rather, both the pharmacological and nonpharmacological treatments seek to reduce the duration, intensity, and frequency of pain in clients (Solomon, 1995). Between the pharmacological and nonpharmacological categories there

are over 30 different treatment options. Treatment choices become even more confusing when turning to research for guidance. A simple perusal of MEDLINE and PSYCHINFO databases reveals that each treatment option has many (from 2 to 12) primary studies touting its "effectiveness." It is apparent in the above description of available treatments that the abundance of choices can be overwhelming for professionals and consumers. The data regarding the effectiveness of all the treatments for migraine headaches needs to be quantitatively analyzed so that treatments can be compared to one another. A meta-analysis is a good vehicle for this kind of analysis of multiple primary studies (Carroll, 1996). Thus, a comprehensive meta-analysis was undertaken to evaluate the effectiveness rates of six empirically supported pharmacological and nonpharmacological treatments for migraine headache syndrome.

Suggested Evidence for Evaluating the Effects of Treatments

It is important to have a common understanding of the meaning of a "potent treatment" before attempting to compare the effects of various treatments. Guidelines have recently been proposed for defining efficacious and effective treatments. Chambless and Hollon (1998) detailed minimum research requirements for an efficacious treatment. First, they stated:

> Treatment efficacy must be demonstrated in controlled research in which it is reasonable to conclude that benefits observed are due to the effects of the treatment and not to chance or confounding factors such as passage of time, the effects of psychological assessment or the

> presence of different types of clients in the various treatment conditions. (p. 7)

The authors also recommend that at least two controlled studies, conducted by independent researchers, need to be conducted for a specified condition. This is a particularly salient issue for many of the pharmacological treatments of migraine headaches. Many medications are recommended for use with migraine headache sufferers, yet lack controlled studies to support their use for this specific purpose.

Chambless and Hollon (1998) described some designs that should be given more weight when attempting to evaluate the efficacy of a treatment. Randomized trials are considered to be the "gold standard" for treatment evaluation and constitute the best evidence. In addition, treatment evaluation should be considered hierarchal in nature. Results that come from treatment studies with no comparison groups should be held more tentatively than results from studies comparing treatment to a control, placebo, or no treatment group. Better yet are results from studies that compare effects of a new treatment to the effects of another established treatment. According to Chambless and Hollon, this is the premier type of efficacy research. Comparison studies that examine the treatment effects of two or more treatments can result in explicit information regarding relative benefits of the two treatments. Unfortunately, in the area of migraine headache there are only a few primary studies that compare the treatments to each other. Those that do compare treatments typically compare them within the same category (e.g., EMG biofeedback to thermal biofeedback or Propranonol with Flunarizine). Primary studies that compare treatments between groups are very rare. It has been suggested that this may

occur because medical and nonmedical professionals become so exclusively focused on their treatment modalities that research collaboration seldom occurs (Holroyd & Penzien, 1990). However, the lack of primary research comparing the numerous treatment modalities available for treating migraine headaches does not abate the need for professionals and clients to be able to evaluate the relative effects of the treatments.

One possible solution to this problem is a large-scale evaluation project. Such a project could compare each available treatment option for migraine headaches, tension headaches, and mixed headaches in a controlled evaluation. However, the number of needed subjects to do a primary research study comparing the success or improvement rates of treatments for migraine headaches would be enormous. Chambless and Hollon (1998) recommended that a minimum of 50 subjects per condition should be used in efficacious research. If this guideline were adhered to, a research study examining outcome of the most common treatments for migraine headaches would require approximately 1700 subjects. Even more subjects would be needed if each treatment group were to be matched with a control group. The project would require collaboration between many groups and professionals and would be costly to conduct. However, the information provided in such a study would be invaluable. Even if this type of large scale study were conducted, replication would be needed before confidence could be gained in the results.

In the absence of such large scale projects another alternative is available. The use of meta-analytic techniques as suggested by Glass (1976) allows for the quantitative analysis of a large number of primary analysis results for the purpose of integrating and

comparing these results. A meta-analytic review comparing all methods of treatments for migraine headaches would provide a way for previously obtained results in each individual area to be compared and evaluated.

Previous Reviews

In preparing to conduct a meta-analytic review, it is appropriate to evaluate previously conducted reviews on related topics (Glass, 1976). This allows the new meta-analytic review to benefit from previously obtained data and avoid pitfalls that others may have encountered. The area of treatment efficacy research for migraine headaches is marked by several well-conducted quantitative reviews and many insightful narrative reviews. Reviews were identified by using combinations of the key words "review," "effective," "treatment," "migraine," "pain," "outcome," and "headaches" in both the MEDLINE and PSYCHINFO databases (1970 to 7/2003). This search strategy identified 102 reviews. Of these, 19 were in a foreign language and 36 presented only information of the pathophysiology or diagnosis of headaches. In addition, a substantial number ($N = 32$) of pharmacological reviews were eliminated because they did not meet selection criteria. For example, many reviews in pharmacological literature present only an author's personal experience with, or opinion of, a treatment for migraine headaches. These reviews do not evaluate controlled studies and provide little information about the effectiveness of the treatments. Thus, the reviews presented here ($N = 12$) met the following criteria: (a) the review provided information regarding the effectiveness of a recurrent headache treatment(s) in either narrative or quantitative form, (b) if the review was narrative then it

must review empirically conducted research, and (c) the review was written or translated into English.

The reviews presented here are evaluated by quality of research design based on six criteria: (a) selecting and delimiting the topic, (b) review of previous reviews in relevant areas, (c) selecting an appropriate sample of studies, (d) data collection procedures, (e) analyzing results, and (f) interpreting and reporting results. Based on these criteria, each review was given an excellent, good, or poor rating. Recommendations based on the literature will be followed by a brief critique of the available reviews in each area.

This paper's examination of previous reviews is broken up into five separate sections. The first examines those reviews which attempt to compare effectiveness rates of treatment(s) from the pharmacological category with treatment(s) from a nonpharmacological category. Again, according to Chambless and Hollon (1998), this type of comparative information is considered to be the "best evidence," thus, these reviews are considered first and in some detail. Then, reviews that considered just pharmacological treatments for headaches are examined, followed by reviews that considered only nonpharmacological treatments. Then, other relevant reviews that provided valuable information about potential correlating variables will be briefly examined. Finally, a recent work that critiques meta-analyses in general will be discussed. A brief summary of major findings and considerations of each existing review can be found in Appendix B.

Reviews Comparing Pharmacological and Nonpharmacological Treatments

There are only two existing reviews that integrate pharmacological and nonpharmacological treatments for recurrent headaches in adults. One of these reviews focused solely on treatments for migraine headaches, the other focuses on tension headaches. The first of these reviews compared the effectiveness of a biofeedback/ relaxation combination treatment to medical administrations of Propranonol in adult migraines. The second review focuses on both pharmacological and nonpharmacological treatments for tension headaches in adults. These meta-analytic reviews are of varying quality, but all provide valuable insight into treatment outcome in recurrent headaches.

Holroyd and Penzien

The first meta-analytic review is an excellent evaluation of the treatments for adult migraine headaches. This review compared the treatment effects of Propranonol and relaxation/biofeedback training. Holroyd and Penzien (1990) used meta-analytic procedures to examine results of primary research studies conducted up to 1989 in migraine suffers (although some mixed headache suffers were included). Four conditions were compared in this review: Propranonol, biofeedback/relaxation, placebo, and no treatment conditions.

Change was variably calculated in the Holroyd and Penzien (1990) review. First, if the primary research articles reported a headache index score, this was used as the change index (the headache index is a composite score that weights and combines intensity and

duration of headaches). However, an undetermined number of studies did not report enough data to calculate the headache index score; therefore, for some studies the authors calculated a "composite headache index score." This score was calculated from an average of reported scores (any or all of frequency, duration, and intensity change scores). Authors reported a significant correlation ($r = 0.86$) between the two methods of calculating change (headache index and their composite headache index score).

Holroyd and Penzien (1990) coded for method of estimating change improvement scores. Some of their primary studies used daily recording journals, where others had practitioners make end-of-treatment gain estimates. The authors found that results differed by approximately 20% if daily headache measures were used to calculate change. That is, the reported improvement in headache activity was 20% lower in studies that had participants rate their headache daily rather than having a therapist or physician rate the change. This difference was significant ($p < 0.001$). Calculations examining only the data from the studies employing the daily recording of headache activity revealed 43.3 % mean improvement in headache in both the Propranonol and the relaxation/biofeedback conditions. Both Propranonol and biofeedback/relaxation showed significantly ($p < 0.001$) higher treatment effects than the placebo and untreated conditions.

These authors based their results on two different types of change scores. In some studies the change score (the headache index) was taken directly from the primary report. In other cases the authors averaged across three variables (duration, frequency, and intensity). The headache index does not account for frequency thus the two methods of computation may have led to dramatically different results. In addition, this review

examined only one type of prophylactic treatment, Propranonol, and excluded the others. Propranonol is the most frequently prescribed prophylactic pharmacological treatment, however, this does not indicate that Propranonol is the most efficacious treatment or the only one worthy of study. Also, Holroyd and Penzien examined only studies through 1989. Over 40 new studies have been published on treatments for migraines in the last 9 years. These additional studies could render some of the Holroyd and Penzien's conclusions outdated. Finally, it is noted that the authors make a particularly important recommendation for future research. They note that while effectiveness rates are compared at the end of treatments it is unclear what long term effects, if any, these treatments have.

Technical Review Conducted by Duke University and the Center for Clinical Health and Policy Research

The second review is a two-part review developed in part for the Agency for Health Care Policy and Research. This study has not been published (likely due to length and complexity) outside of the NTIS Website. The review was posted (at NTIS) after the proposal date of the current project and it mirrors many of the current author's intentions. The first part of the part of this review (authored by Duke University and The Center for Clinical Health Policy and Research, 1999a) synthesizes the data for behavioral and physical treatments for migraine. The second part (Duke University and The Center for Clinical Health Policy and Research, 1999b) synthesizes the data for prophylactic pharmacological agents for migraine. This methodologically sophisticated review is

excellent in most aspects. The review only included randomized controlled trials of pharmacological, psychological, behavioral and physical (e.g., acupuncture) treatments. A comprehensive group of findings are presented that include more than the six treatments evaluated by the current meta-analyses. Findings from this excellent quantitative review indicated that most pharmacological and nonpharmacological treatments had functionally equivalent effect sizes (all of which were in the moderate range). A similar pattern and magnitude of findings was reported for nonpharmacological treatments.

Standardized mean difference effect sizes for nonpharmacological treatments ranged from .08-1.61. The nonpharmacological treatments and summary weighted mean effect sizes are as follows: relaxation (*es* = .55), thermal biofeedback (*es* = .38), thermal biofeedback plus relaxation (*es* = .40), E.G. biofeedback (*es* = .77), cognitive behavioral therapy (*es* = .54), and thermal biofeedback plus cognitive therapy (*es* - .37). The pharmacological meta-analysis examined 52 treatments that included hormonal treatments for menstrual migraine. They report that the top seven agents with their effect sizes are naproxen sodium (*es* = .62), Flunarizine (*es* = .52), Propranolol (*es* = .55), amitriptyline (*es* = .62), Timolol (*es* = .69), pizotifen (*es* = .91), Divalproex Sodium (*es* = .93). The other pharmacological agents examined in this study were reported to have variable results with low effect sizes.

However, this meta-analysis does not compare the findings between the two groups. Instead they report their findings in two separate reports in two separate formats. Of additional interest was the authors' choice in this review to use varied statistical procedures for pharmacological studies and nonpharmacological studies. Standardized

mean difference scores and odds ratios are calculated for pharmacological studies, non pharmacological studies include standardized mean differences and percent improvement measures. This is curious given the emphasis on consistency of evaluation methods across domains. The percent improvement scores were used in the Duke study as a cross check for standardized mean difference scores. The authors reported that the percent improvement scores in some cases underestimated treatment effects. In addition, percent improvement scores have been criticized because this procedure averaged unstandardized data and can be misleading.

Summary of Reviews That Compare Pharmacological and Nonpharmacological Treatments

A total of two reviews exist that evaluate pharmacological and nonpharmacological treatments for recurrent headaches. An evaluation of the these reviews indicates that: (a) Propanonol, relaxation training, biofeedback, and cognitive therapies all have been reported to produce similar treatment outcomes and improvements in migraine sufferers (both adults and children); (b) percent improvement scores may underestimate treatment effect; and (c) type of outcome measure highly influences the results of the study.

Cumulatively, the reviews identify several pitfalls for future researchers. First, change scores should be calculated for frequency, intensity, and duration if headache scores are not available. If standard deviations are presented in primary literature, standardized effect sizes should be calculated. Next, it may be wise to use all available studies and code for quality rather than eliminate studies due to methodological flaws.

Finally, when examining both pharmacological and nonpharmacological studies researchers should report specific treatment effects, in addition to categorical effects (i.e., pharmacological or nonpharmacological).

Pharmacological Reviews

This section examines those reviews that focus on pharmacological treatments. Two quantitative reviews on migraine headaches exist in this area. (Holroyd, Penzien, & Cordingley, 1991; Onghena & Van Houdenhove, 1992). Other narrative reviews are available (e.g., Capobianco et al., 1996; Solomon, 1995), however, the reviews are primarily informational in nature and focus on side effects of the medications and contraindications. This section will focus only on those reviews that are quantitative in nature.

Holroyd, Penzien, and Cordingley

The Holroyd et al. (1991) review examines existing literature on the effectiveness of Propranonol as a prophylactic treatment for recurrent migraines. This review synthesized data from 53 studies, which resulted in a total of 73 groups (Propranonol and placebo). The Holroyd et al. review calculated two different types of change scores. The first was pre-post treatment gains, the second subtracted gains of the placebo groups from the gains of the treatment (Propranonol) groups. Change scores were analyzed separately and results varied significantly by method of calculation. That is, change scores were lower when placebo groups were used in the calculation instead of pretreatment scores.

The Holroyd et al. review concludes, that on, average prophylactic use of Propranonol results in a 44% reduction in migraine activity when the most conservative outcome measures are used (patient daily diaries) and when change is compared to placebo groups. Estimates are higher when less conservative measures are used (clinician ratings of pre- and posttreatment scores).

This review is considered to be excellent because it has clear selection criteria and identifies correlates of treatment outcome (age, gender, mortality, mortality due to side-effects, dose of active agent, chronicity of headache, and quality of study). While providing excellent information about the use of Propranonol, the net result is information that is difficult to interpret. This is because Propranonol is compared only to placebo groups and not other treatments. Thus, it is difficult to identify whether Propranonol is better, equal to, or worse than other prophylactic treatments. The authors of this review report that Propranonol has equal effectiveness rates to other pharmacological prophylactic treatments, but no data are provided to support this position.

Onghena and Van Houdenhove

Onghena and Van Houdenhove (1992) conducted a good meta-analytic review of antidepressant-induced analgesia in chronic pain patients. Thirty-nine placebo controlled studies were examined; 10 were conducted with migraine, tension headache, or mixed headache subjects. This review provides an adequate description of selection procedures, the problems with primary research validity, quality of study, compliance, and types of medication used. Conclusions indicate that patients with migraines, tension headaches, and

mixed headaches showed statistically significant improvements in their headache activity when using an antidepressant medication. In addition, larger effect sizes were noted for recurrent headache patients than any other pain category.

Yet, the findings in the Onghena and Van Houdenhove (1992) review are difficult to interpret because they do not compare the analgesic effects of antidepressants to analgesic effects of any other pharmaceutical agent. The results of this meta-analysis are important, but difficult to integrate with other analgesic findings. However, this review is one of the few that quantitatively evaluate pharmacological treatments, thus providing important design information. In this study it was recorded whether other drugs were allowed to be taken during the study, the activating profile of the drug (psychomotor activation, nonactivating/nonsedating, and sedating), in-patient status, mean age, duration of pain, side effect profiles of the drugs, numbers of subjects dropping out due to side effects, and patient profiles indicating drug selection.

The last point is of particular interest. Many of the pharmacological interventions have a highly selective process to determine if subjects can safely take the medication. Selection processes may include the consideration of pre-existing medical conditions, age, previous response to medications, current medications taken, and side effect profiles of the pharmacological agent (Solomon, 1995). This selection process may introduce a selection bias that is not comparably matched in the nonpharmacological research that may cause problems in the comparison of the two interventions. To best attempt to compare the pharmacological and nonpharmacological treatments the above information should be gathered on all primary studies (pharmacological and nonpharmacological).

Summary of Pharmacological Reviews

Both of these reviews offer insight into procedures that are helpful when reviewing pharmacological literature. For example, Holroyd et al. (1991) suggested that change scores should be calculated in two forms and then results should be analyzed separately. The Onghena and Van Houdenhove (1992) review suggests that reviews of pharmacological research should examine several variables. These include; identification of other medications taken (analgesic and others), inpatient status, age, and duration of pain. In addition, both reviews (Holroyd et al.; Onghena & Van Houdenhove) indicated that overall mortality rates versus mortality rates due to medication side effects should be coded separately.

Nonpharmacological Reviews

One review that examines adult migraine headache treatment is available in the nonpharmacological area. Compas, Haaga, Keefe, Leitenberg, and Williams (1998) examine the outcome of biofeedback, relaxation, and cognitive treatments. Unlike the pharmacological reviews, this review makes comparisons between treatments that provide valuable insight into the effectiveness of treatments in relation to one another.

In narrative form, Compas et al. (1998) reviewed all empirically supported nonpharmacological treatments for smoking, cancer, chronic pain, and bulimia nervosa. A subsection is included on migraine headaches that reviews thermal and E.G. biofeedback, relaxation, cognitive therapy, and cognitive behavioral therapy. Most of the information in this narrative review is based on previously mentioned quantitative reviews. The authors

report that many of the previously conducted meta-analyses have been supported by more recent work. Nevertheless, there was no attempt to quantitatively review the more recent work. Furthermore, Compas et al. evaluated the effectiveness of cognitive and cognitive-behavioral treatments (both of which are treatments that have not been adequately quantitatively reviewed) and concluded that cognitive treatments are not effective in reducing migraine headache activity.

The Compas et al. (1998) review was written simultaneously with the Chambless and Hollon (1998) recommendations for evaluating empirically supported treatments. As such, Compas et al. clearly identify studies that are randomized, provide correlational evidence, and compare to other treatments. The list of studies generated by Compas et. al, while not inclusive, would be essential to include in a future meta-analysis.

Other Relevant Reviews

Four other reviews will be briefly discussed. They are not specifically focused on the treatment of migraine headaches, but they provide valuable information about potential correlates and coding procedures. Because these reviews are less directly related to the current topic, they will only be discussed briefly.

Flor, Fydrich, and Turk (1992)

This is an excellent meta-analytic review which investigated the efficacy of multidisciplinary pain clinics. The authors investigated age, marital status, education level, socioeconomic status (SES), employment status, compensation, litigation, and medication

use. However, none of the above were found to correlate significantly with study outcome. Though the Flor et al. study focused primarily on the treatment of back pain patients in pain clinics, the noted correlates may be important when evaluating migraine headaches.

Lander

Lander (1990) also provided a narrative review on pain management. While not specifically addressing migraine headaches, Lander highlights the importance of client and researcher variables in evaluating pain management techniques. Variables considered to be important are: (a) practitioner inference about treatment gains "polluting" the outcome measures, (b) potential differences in convenient, solicited, clinical, and nonclinical populations, (c) gender, SES, and age as mediating variables, and (d) duration of pain.

Holroyd and Penzien

In 1986, Holroyd and Penzien conducted a meta-analysis that examined both client variables and treatment variables that may impact the treatment of tension headache. This review focused solely on tension headaches and nonpharmacological interventions. Type of treatment, length of treatment, training of therapist, and attention to transfer training did not significantly correlate with treatment outcome. However, sample characteristics were found to significantly correlate with the reported reduction in headache activity. Most dramatically, treatment outcome varied by age, gender, and solicited, rather than referred clients. Primary studies with younger mean ages, more females, and solicited samples all reported stronger treatment effects.

Malone and Strube

In contrast to Holroyd and Penzein (1985), Malone and Strube (1988) found that sample characteristics were not correlated with outcome of psychological treatments. Malone and Strube conducted a meta-analysis of effective nonpharmacological treatments. However, Malone and Strube collapsed across types of pain treatment (e.g., headache, backache, cancer). In addition, Malone and Strube calculated change scores for primary studies that did not report pretesting by estimating from studies that did have similar pretests. This procedure is based on questionable assumptions and is likely to mask true differences. Given the procedural errors, the Malone and Strube meta-analysis was given a poor quality rating and thus, the evidence presented in the Malone and Strube review is not convincing.

Reviews of Meta-Analytic Reviews

Finally, one excellent quantitative review has been conducted using previous meta-analytic reviews that have examined analgesic interventions (Jadad & McQuay, 1996). The article is primarily focused on identifying weaknesses in the current reviews and making recommendations for reducing bias and improving quality. Jadad and McQuay report that a simple MEDLINE search missed nearly 50% of relevant articles making many of the medically based meta-analyses incomplete. In addition, the importance of completely identifying the methods used to locate primary studies and describing criteria for assessing validity of primary studies was highlighted. While these recommendations are standard for conducting a meta-analysis, Jadad and McQuay report that 26% of meta-analyses

reviewed did not identify the selection methods, and 21% did not identify design characteristics of their sample.

Other important issues in meta-analysis research identified by Jadad and McQuay (1996) are the quality and validity of the primary study. As mentioned above, many authors have reported that quality of primary study is an important outcome variable (e.g., Blanchard, Andrasik, Ahles, Teders, & O'Keefe, 1980; Bogaards & ter Juile, 1994; Holroyd & Penzien, 1985; Onghena & Van Houdenhove, 1992). Authors of meta-analyses often evaluate the quality and validity of primary studies because it has been hypothesized that primary studies of poor design with threats to validity may produce higher effect sizes than studies of good quality (Slavin, 1984). Therefore, the conclusions of a meta-analysis may be overestimates if the author does not account for the quality of the primary studies it includes. Jadad and McQuay reported that while many authors report that they evaluated quality of study, 60% of meta-analyses did not describe methods of assessing primary study validity and quality. Jadad and McQuay reported that a good meta-analysis includes a systematic and objective method of assessing quality and validity of primary studies that is specifically described in the text.

In summation, there are 12 applicable reviews in this area. Only a few of these reviews have compared treatments to each other. Two compared treatments across categories (nonpharmacological and pharmacological), but only one of these reviews (Holroyd & Penzien, 1990) focused directly on migraine. The Holroyd and Penzien review did not comprehensively examine multiple pharmacological and nonpharmacological agents. A review of multiple pharmacological and nonpharmacological treatments is

available for tension-type headaches (Bogaards & ter Kuile, 1994), but none exist in the migraine literature. The remaining reviews studied the outcome of treatments within a single category (i.e., relaxation and biofeedback in the nonpharmacological category). The pharmacological reviews did not compare treatments at all, instead, they compared treatments only to placebo or control groups. The net result of these reviews is a wealth of information that is almost as difficult to interpret and use as the individual primary studies.

Summary

Migraine headaches are a disabling problem for millions of Americans. There are over 30 treatments available for migraine headaches. Much of the primary research in the area of migraine headaches is focused on identifying the effectiveness rates of these treatments in comparison to control or placebo groups. This procedure does not allow for the comparison between treatments. Only seven primary studies have conducted research that compares a pharmacological and nonpharmacological intervention. As a result, there is a wealth of data on individual treatments that is difficult to compare to one another and use (Chambless & Hollon, 1998). A synthesis is needed to make the data more manageable. There have been a variety of systematic reviews conducted in this area, however, their focus is too narrow and some are now outdated. Therefore, a more comprehensive meta-analysis that includes an evaluation of both pharmacological and nonpharmacological treatments on migraine headaches needs to be conducted.

CHAPTER III

PURPOSE OF STUDY AND RESEARCH QUESTIONS

The purpose of this study was to integrate data available on pharmacological and nonpharmacological treatments for migraine and mixed headaches. Mixed headaches were being included because of the high overlap between the two syndromes. This study is a meta-analysis and, thus, will quantitatively compare effectiveness rates of many different treatments.

1. How do the effectiveness rates of treatments compare to each other and between groups (pharmacological and nonpharmacological)?

2. What sample characteristics are significantly correlated with outcomes? Do these characteristics vary by treatment category?

3. Do outcomes vary if broken into two levels of empirical evidence (placebo control group comparison, and multiple treatment comparison) as suggested by Chambless and Hollon (1988)?

4. What information is available regarding short and long term effectiveness rates for pharmacological and nonpharmacological studies?

CHAPTER IV

METHODS

The methods used in this meta-analysis are described below. First, procedures for selecting the treatments and primary articles to be evaluated are discussed. Then, a description of the sample with a list of articles included in the meta-analysis is provided. Data collection and quality rating methods are discussed next. Finally, methods and rationales for analysis are presented.

Selection of Sample

The selection of studies to be included in the current meta-analysis followed a multistep procedure. First, popular classes of treatment were identified, then the most frequently researched treatments within the class were identified. Once identified, all outcome studies regarding the treatments were gathered. A search strategy was developed to ensure comprehensiveness. Once primary articles were obtained, inclusion and exclusion rules were applied. Each of these steps are described in turn below.

The current meta-analysis identified six different treatments for migraine prophylaxis to compare. Treatments to be evaluated were selected by conducting a preliminary literature search of available classes of treatments. In the pharmacological treatment area a preliminary literature search identified three major pharmacologic classes that are currently being researched for use as prophylactic agents: beta-blockers, calcium agonists, and anticonvulsants. Once these classes of medications were identified, a more

comprehensive abstract search identified which pharmacological agent had the greatest number of controlled published research outcome studies. This process identified the top three agents in each category; Propranonol, Flunarizine, and Divalproex Sodium. This process was duplicated for nonpharmacological prophylactic treatments for migraine. The preliminary literature search in the area of nonpharmacological treatments identified thermal biofeedback, relaxation therapies, and mixed therapy treatments as the most researched areas.

Once the treatments for inclusion in the meta-analysis had been identified, a more comprehensive literature search began. A search for published, peer-reviewed literature on the effectiveness of each treatment was conducted. Literature published between 1970 to July, 2003 was included in the search. This search included searches of MEDLINE, PSYCHINFO, ERIC, and the Current Contents databases. Key words used in database searches included combinations of the following: the treatment names, controlled, random, effectiveness, prophylaxis, migraine, headache, vascular headache, treatment, and outcome. In addition, reference sections of reviews on these treatments were checked and cross-referenced to ensure comprehensiveness. Finally, the last year (2002-2003) of *Pain, Headache, Journal of Consulting and Clinical Psychology, Journal of the American Medical Association, & Cephalalgia* were reviewed for newly published articles that may have not been accounted for in the computer-based search.

After articles had been identified in the above-mentioned treatment categories, a set of inclusionary and exclusionary rules were applied to each article. Inclusion rules included (a) three studies examining a single treatment for migraine headaches must exist

for a treatment to be considered in the meta-analysis, (b) studies had to have at least five subjects in each condition to be considered in the meta-analysis, (c) study evaluated a pharmacological or nonpharmacological treatment for migraine or mixed headache, and (d) study must provide enough numerical detail to calculate at least one change score per condition. Studies were excluded from the analysis if (a) the study did not compare the treatment to a placebo, control group or comparison treatment,(b) the study was conducted with animals, (c) the study only examined children under the age of 18, (d) the study was published in a foreign language, (e) the study reported on the same data that had already been reported elsewhere, (f) the study employed a single-subject research design, and (g) the study was an analogue study.

Description of Sample

Approximately 400 articles were obtained and cataloged via a database system. This number includes approximately 200 articles that were eliminated because they evaluated the wrong treatments, were reviews, did not provide original data, provided practitioner guidelines only, or had a focus other than treatment outcome evaluation. The selection criteria were then applied to the 191 remaining articles. Careful documentation was kept on each article that was excluded. This process resulted in examination of 57 articles in the Propranonol category and the retention of 30 articles, 52 articles in the Flunarizine category and the retention of 17 articles, and 18 articles in the Divalproic Sodium category and the retention of seven articles. In the nonpharmacological area many of the primary studies were multi-arm studies that provided information on more than one

treatment being included in the analysis. Thus, the sample is best described as a sample of 64 articles from which 15 studies were used to calculate 18 effect sizes in the thermal biofeedback category. Four articles were retained in the relaxation therapy category. Nine articles were used in the mixed therapy category (five EMG biofeedback articles and four thermal biofeedback plus relaxation plus cognitive therapy). This resulted in 82 studies being retained overall and included in the study, and 109 studies being excluded.

It should be noted that a large number of nonpharmacological studies were excluded due to insufficient reporting of data in the write up. Many studies in the nonpharmacological area were well-designed controlled studies. Interestingly, a number of authors attempted to use forms of relaxation or meditation as a control condition. However, these "control groups" have been identified as active treatments within the primary study and in subsequent research. Thus, the studies using relaxation controls are actually comparison studies between two treatments. Because of their conceptual premise, commonly authors had developed a priori hypotheses only about the difference between the so-called control group and the treatment group. Therefore, many of these studies only reported a nonsignificant finding for an independent *t* test. Nonsignificant results can be used to calculate an effect size if exact *t* and or *p* values are given. However, the common practice for reporting non-significant results is with lower bounds ($p > .05$) or without numerical data (t = N.S.) Reporting results in this manner precludes the calculation of effect sizes, and thus violates inclusionary rule (d). This is an unfortunate loss of data and results in a relatively small number of articles being available for

evaluation in the nonpharmacological treatment area. Those included are listed below by categories in Tables 1 to 4. Lists of excluded can be found in Appendix C.

Data Collection

Once a study was identified as meeting all of the criteria for being included in the current meta-analysis, it was coded for a variety of information (see Appendix D for copy). Variables coded included sample used in the primary study, sample size, year the study was published, percent female, type of control group used, randomized assignment, mortality rate, and design of study (parallel or cross-over). In addition, quality of study was coded in two separate ways. The first quality of study measure used is a research-based instrument developed by Jadad et al. (1996). This instrument codes randomization, double-blind features, and descriptions of dropouts. It been shown to have high inter-rater reliability because of its objective descriptions of criteria (Jadad et al.). Scores for articles being rated range from zero to five, with higher scores indicating a better quality report. The second measure was created to examine quality issues specifically related to the headache literature. Items include presence of wash-out or run in periods, types of recording used, comparison groups used, type of setting, and control of medications/ treatments. This measure has scores ranging from one to three, with higher scores indicating a better quality of report (see Appendix E for copy). This measure was checked for inter-rater reliability by having a second PhD level psychologist code a randomly chosen set of articles (approximately 15% of the total). The reliability coefficient for the

Table 1

Primary Articles That Were Included in the Meta-Analysis for Propranolol

Authors
Ahuja & Verma (1985)
Albers, Simon, Hamik, & Peroutka (1989)
Andersson & Petersen (1981)
Borgesen, Nielsen, & Moller (1974)
Daholf (1987)
Diamond & Medina (1976)
Diener et al. (1996)
Forssman, Henriksson, Johannsson, Lindval, & Lundin (1976)
Havanka-Kanniainen, Hokkanen, & Myllyla (1988)
Johnson, Hornabrook, & Lambie (1986)
Kangasniemi & Hedman (1984)
Kangasniemi, Nyrke, Lang, & Petersen (1983)
Kjaersgard Rassmussen, Holt Larsen, Borg, Soelberg Sorensen, & Hansen (1994)
Kuritzky & Hering (1987)
Lucking, Oestreich, Schmidt, & Soyka (1988)
Ludin (1989)
Mathew (1981)
Mikkelsen, Kjaersgaard Pedersen, & Christiansen (1986)
Nadelmann, Phil, Stevens, & Saper (1986)
Pita, Higueras, Bolanos, Perez, & Mundo (1977)
Pradalier et al. (1989)
Rao, Das, Taraknath, & Sarma (2000)
Rosen (1983)
Sargent et al. (1985)
Standness (1982)
Stensrud & Sjaastad (1976)
Tfelt-Hansen, Standnes, Kangasneimi, Hakkarainen & Olesen (1984)
Weber & Reinmuth (1972)
Wilderoe & Vigander (1974)
Zeigler, Hurwitz, Preskorn, Hassanein, & Seim (1993)

Table 2

Primary Articles That Were Included in the Meta-Analysis for Flunarizine

Authors
Al Deeb, Biary, Bahou, Al Jabeeri, & Khoja (1992)
Allias et al. (2002)
Bussone et al. (1987)
Freitag, Diamond, & Diamond (1991)
Frenken & Nuijten (1984)
Lamsudin & Sadjimin (1993)
Louis (1981)
Louis & Spierings (1982)
Lucking, Oestreich, Schmidt, & Soyka (1988)
Ludin (1989)
Mentenopolous, Manafi, Logothetis, & Bostantzoulou (1985)
Nuti et al. (1996)
Pini, Ferrari, Guidetti, Galetti, & Sternieri (1985)
Rascol, Montastruc, & Rascol (1985)
Sorensen, Hansen, & Olesen (1986)
Sorensen & The Danish Migraine Study Group (1989)
Thomas, Behari, & Ahuja (1991)

Table 3

Primary Articles That Were Included in the Meta-Analysis for Divalproex Sodium Studies

Authors
Jensen, Brink, & Olesen (1994)
Kaniecki (1997)
Kinze et al. (2001)
Klapper (1997)
Lenaerts, Bastings, Sianard, & Schoenen (1996)
Mathew et al. (1995)
Rothrock, Kelly, Brody, & Golbeck (1994)

Table 4

Primary Articles that Were Included in the Meta-Analysis on Nonpharmacological Treatments

Author	Thermal biofeedback	Relaxation therapy	Combined therapies	
			EMG	TBF+
Blanchard, Andrasik, Neff, Arena, et al. (1982)	X	X		
Blanchard et al. (1985a)	X			
Blanchard et al. (1985b)	X			
Blanchard, Appelbaum, Nicholson, et al. (1990)	X			X
Blanchard, Appelbaum, Radnitz, et al. (1990)	X	X		X
Blanchard, Nicholson, et al. (1991)	X			
Blanchard et al. (1997)	X			
Blanchard et al. (1978)	X	X		
Daly, Donn, Galliher, & Zimmerman (1983)	X	X	X	
Gauthier, Lacroix, Cote, Doyon, & Drolet (1985)	X			
Gauthier, Cote, & French (1994)	X			
Holroyd et al. (1995)	X			
Holroyd et al. (1988)	X			
Jurish et al. (1983)	X			

(*table continues*)

Author	Thermal biofeedback	Relaxation therapy	Combined therapies	
			EMG	TBF+
Lake, Raney, & Papsdorf (1979)	X		X	X
Largen, Mathew, Dobbins, & Claghorn (1981)			X	
McGrady, Wauquier, McNiel, & Gerard (1994)			X	
Mullinix, Norton, Hack, & Fishman (1978)	X			
Wauquier, McGrady, Aloe, Klauser, & Collins (1995)			X	

overall score was .88, indicating that a high level of agreement between raters. This suggests that this measure can be reliably used by different raters.

Analysis

After data were collected from primary studies, a variety of procedures were used to evaluate the data for use in this meta-analysis. Analysis procedures will be described in the following order: (a) standard effect sizes for continuous data, (b) odds ratios for binary data, (c) inclusion of cross-over studies for continuous and binary data, (d) homogeneity model, and (e) inferential statistics used.

Effect Sizes for Articles Reporting Continuous Data

Data from primary articles that are expressed in the form of continuous outcomes

use means, standard deviations, and inferential statistics to report the outcomes of the study. When this is the case the standardized mean differences were calculated. Glass, McGaw, and Smith (1981) report procedures for this calculation in its simplest form as g = (M treatment) - (M control)/SD pooled. Hedges and Olkin (1985) reported that this formula results in a small sample bias and, thus, this meta-analysis used d as the estimator of effect size ($d = g(1 - 3/4m - 1)$), where m equals the degrees of freedom based on the pooled standard deviation calculation. Throughout this meta-analysis, standard procedures for estimating effect sizes were used as recommended by Hedges and Olkin. All standard deviations used in calculations are pooled.

A key statistical issue in examining continuous data on migraine prophylaxis is determining the appropriate statistical formula to use when paired data are presented in the primary article. Most research on migraine prophylactics use a pre-post design, even if a control/placebo group is being used. Formulas exist for calculating a paired t effect size given the N, SD and paired t-statistic. However, this type of data (i.e., the N, SD *and* paired t-statistic) is rarely presented in peer-reviewed journal articles. Dunlop, Cortina, Vaslow and Burke (1996) convincingly argued that the calculation for the standardized mean difference using the independent t-test formula value can be substituted in lieu of the paired-data formula. They argued that if the paired t test is used (which corrects for the amount of correlation between the measures), the resulting ES estimate will be an overestimate of the actual ES. Therefore, this meta-analysis employed the more conservative independent t-test method when presented with paired data.

Given the variety of statistical information provided in the primary articles reporting continuous data, a hierarchy was developed to aid in consistent decision making about statistical selection. Examples of all the below calculations can be found in Appendix F. All continuous data calculations were performed with the aide of Meta-Stat ® computer program.

Decision Making About Effect Size Estimate

1. The "premiere" statistic is the standardized mean difference between two independent samples.

2. If data to calculate the above is not provided, then an effect size estimate from *t, F, Z,* or *r* will be used.

3. If 1 or 2 cannot be calculated, then an exact *p* estimate will be used to calculate a *T* score.

4. If the groups being compared are a baseline average of several treatments, then a pooled standard deviation will be calculated and the groups will be compared as if independent.

5. A *z*-score conversion of percent improved will be calculated if appropriate data are provided. A *z*-score conversion was only calculated on single-population studies if there were three discrete assessment periods (a baseline, a placebo, and an active treatment). Odds ratios are to be calculated in preference to *z*-score conversion if data are present.

Effect Sizes for Articles Reporting Binary Data

The second type of data available in primary articles on migraine prophylaxis is binary data. Medical research in the area of migraine headache commonly divides outcome data into binary form. That is, researchers often report data for "responders" (those who achieved a 50% reduction on some measure through treatment) and "nonresponders" (those who achieved less than 50% reduction on some measure through treatment). This is valuable data primarily because individuals are used as their own control, which lends itself to a broader range of statistical analysis. The reduction of headache indices at the 50% level is considered to be a clinically significant change and is used industrywide. The procedure of reporting responder results also has draw backs. Namely, the data are artificially dichotomized which results in the loss of exact data in some cases. That is, those subjects who respond at the 48% level are treated the same as those who respond at the 2% level. Likewise, those who respond at the 55% level are treated the same as those who respond at the 100% level. Thus, binary data is less exact than continuous data and carries less interpretive weight than continuous data. In the present study, more confidence is placed in results obtained from continuous data (because they are considered to be more representative of the primary findings) than in results gained from the binary effect sizes.

Given that many studies initially selected for inclusion in this meta-analysis reported binary data, a standardization statistic is needed so that it can be compared to other studies. There are three alternative measures that are generally considered for

standardizing binary outcome: the odds ratio, the risk difference, and the risk ratio. The three statistics vary based on the measurement and use of absolute size difference and relative size difference of treatment effect. The current meta-analysis reports binary data in the form of odds ratio's (OR). Deeks (2002) reported that odds ratios are typically used in case-control studies when disease prevalence is not known. In the current research the baseline prevalence is unknown (i.e., the placebo or control rate of spontaneous remission). Deeks reported that the use of relative risk ratios or risk difference ratios, while potentially more intuitive, alters the effect measure by entering a false prevalence rate that is not known and would "obviously be wrong, so odds ratios are the ideal" (p. 1598). The OR has a neutral statistic of "1," which indicates the odds of receiving benefit/harm from the treatment is equal in both the control and the treatment group. As the numbers approach infinity or "0," they indicate what the odds of benefit/harm are for one group in comparison to the other. Thus, a .73 OR indicates that you have .73 times higher chance of getting benefit/harm in the first treatment than in the second. Values over one simply indicate that the benefit/harm chance is higher in the second group than it is in the first. If confidence intervals encompass the value of 1, then the chance of receiving benefit/harm from the two treatments are considered to be nearly equal. When researchers report OR they typically maintain one treatment group as the "anchor," in this study the control was always the second treatment group, indicating that higher scores favor the active treatment.

Once OR were calculated they were converted to Log-odds or the natural logarithm of the OR. This converts the metric from an asymmetrical distribution that is

distributed between zero and infinity to a symmetric statistic running from minus infinity to plus infinity, with zero being the neutral value. This makes it easier to compare negative with positive associations. And it is necessary in order to aggregate a total effect (Deeks, 2002). All OR were calculated with the assistance of MetaAnalysis 3.0 ® by Alan Chang 2001. Examples of an OR calculation can be found in Appendix F.

Inclusion of Cross-Over Design Studies

In the past many authors of meta-analytic studies have chosen to disregard studies that employ a cross-over design. A preliminary review indicated that 82% of meta-analyses at the Cochrane Controlled Trials Registrar do not include cross-over design studies (Elbourne, Altman, Higgins, et al., 2002). This practice has been largely founded on poor reporting of data in cross-over trials. Many primary studies choose to report grand mean and standard deviations for treatment groups. This disregards the fact that patients are their own controls, each patient has received multiple treatments, and means are actually change scores and not independent group means. Thus, many conducting meta-analyses have chosen to ignore data from cross-over trials.

Excluding cross-over data has come under fire recently due to the exclusion of large numbers of studies that compare treatments (e.g., Curtin, Altman, & Elbourne, 2002; Elbourne et al., 2002). Theoretically, cross-over data is more robust than parallel data. In addition, cross-over designs often compare two active treatments to each other. Chambless and Hollon (1998) suggested studies that compare a new treatment to an established treatment can be considered the premier type of efficacy research. Thus, the

exclusion of cross-over design studies would lead to the exclusion of studies that are both statistically and theoretically important. Consequently, it was determined that the present study would include cross-over designs.

How cross-over designs should be included in a meta-analysis is a matter of statistical debate. Three general options are available: (a) use pooled data from cross-over studies and treat them as if the datum came from a parallel design (which results in an underestimation of effect); (b) use data only from one phase of the cross-over design thus making the study a parallel one (which results in an inflated Type 1 error); (c) include only studies that report individual change scores, calculate the pooled variance estimate, and calculate paired statistics effect sizes to adjust for the cross-over design (results in the loss of a large number of studies).

The current meta-analysis employed a combined approach described in Elbourne et al. (2002) for articles reporting continuous data. Primary studies that listed data on individual patients or accurately provided paired data analysis were used to generate Pearson's *r* statistics for between group correlations (accounting for the fact that in a cross-over design individuals are their own controls). The lowest correlation found among primary articles was then substituted into the effect size calculation for the other cross-over design studies within the same treatment type. An example of the substitution of a correlation into a calculation for a standardized mean difference effect size can be found in Appendix F.

Odds ratios require a different strategy when cross-over studies are included. Given that OR already take into account the odds for a single individual on multiple

treatments, correlations adjustments have not been found to alter the estimate of the treatment odds (Elbourne et al., 2002). Articles reporting binary data that were summed across trial periods were used as if they were parallel.

Homogeneity Model

The selection of the appropriate summary statistic that represents the average treatment effect size and corresponding confidence intervals for the meta-analysis is based primarily on the issue of variance. Standard procedures for analyzing results of meta-analyses have traditionally been based on the assumption that there is one true treatment effect being described by the meta-analysis (Fixed Effects Model). Under the Fixed Effects Model (FEM) there is an assumption that any variation that occurs among study effect sizes is caused by study variance accounted for by sampling error, patient characteristics, or study characteristics (i.e., variance with-in the study). Summary statistics based on this assumption assume that each study in the meta-analysis has the same underlying effect (Brockwell & Gordon, 2001). That is, the FEM assumes that there is little study-to-study variation and that the selection of studies is homogeneous.

In contrast the Random Effects Model (REM) assumes that the studies represent a heterogeneous sample with two sources of variance--the variance within the study and the variance between the studies. The addition of the between-study variation to the model indicates that the primary studies each contribute (potentially uniquely) to the true effect. Thus, the summary statistic becomes an indication of the general treatment effect that centers around the true effect (Brockwell & Gordon, 2001; Glasziou & Sanders, 2002).

The two separate sources of variance allow for the assumption that the primary article effect sizes are independent and normally distributed (Brockwell & Gordon). The REM introduces an estimation of between study variance into all weighted averages to account for a random model.

The choice between the two models can be made by testing for heterogeneity between the individual effect sizes. The amount of heterogeneity between studies can be tested through the use of the statistic defined by Cochran (1937), which in practical form is:

$$Q_{\hat{w}} = \sum \hat{w}_i (Y_i - \hat{\mu})^2 .$$

Where w = the inverse of the variance (or weight), Yi = the effect of each trial, and $\hat{\mu}$ = the overall effect estimated from the meta- analysis. Thus, the calculation sums together the weighted differences between overall effect size and the individual study effect sizes (for an example of a calculation of a Cochran's Q see Appendix F). As the value of Q_w increases (results can range from zero to infinity), it indicates more study to study variation. Cochran's Q is typically reported with a p value to indicate if heterogeneity is considered to be more than would be expected by random sampling error. Thus, a significant p value indicates that more heterogeneity exists than would be expected by chance. The current study calculated Cochrane's Q for each treatment type. However, it has been argued that the choice between FEM and REM should not be solely based on the Cochran's Q (Brockwell & Gordan; Chang, Waternaux, & Lipsitz, 2001). In addition, because this meta-analysis is comparing six separate treatments, it is important for a single

model to be used consistently across treatment types. Thus, the current meta-analysis applied a REM to all summary statistics. The FEM has been shown to have too narrow of confidence intervals and to over-estimate the "treatment effect" being described by the meta-analyses. Larger confidence intervals and smaller effect sizes are likely to be produced when using this model than when using the FEM (Brockwell & Gordan, 2001; Curtin et al., 2002). The calculation of the Cochrans *Q* and summary statistics in the REM model were calculated with the assistance of MetaAnalysis 3.0 ® computer program developed by Alan Chang 2001.

Inferential Statistics

Rosenthal's file drawer calculation was calculated for each treatment study. The file drawer phenomenon refers to a well-known publication bias toward significant results. That is, publishers often only publish articles that report significant findings. Articles that do not report significant findings are often unpublished and left in the author's "file drawer." Meta-analyses that only consider published articles are then basing results on a biased sample. The Rosenthal's file drawer calculation is considered a rough guide for determining the number of unpublished null findings that would be needed to threaten the findings of the meta-analysis on the published articles. Essentially, the Rosenthal's file drawer test indicates if a finding of a meta-analysis is robust enough to stand against null unpublished results. Each calculation is reported with the following: (a) an *N*, that represents the number of null studies it would take to threaten the finding;, (b) a "credible if" statement indicating that if the *N* is lower than the "credible if" number, then the

findings may not be robust to the file drawer phenomenon; and (c) a probability level based on a *z* score indicating the probability of finding a similar result by chance.

Regression analysis was used by treatment to identify study characteristics that correlated with treatment effects. Pearson product-moment correlations were run to identify significant associations between the study or sample characteristic and outcome. Comparison of summary statistics across treatments through both visual comparison and analysis of variance (ANOVA) were conducted to test for statistical differences between summary statistics.

CHAPTER V

RESULTS

The results of the meta-analysis will be presented according to the following sections: (a) results for Propranonol, (b) results for Flunarizine, (c) results for Divalproex Sodium, (d) results for thermal biofeedback, (e) results for relaxation treatments, (f) results for combination treatments, and (g) comparisons between summary statistics for all treatments.

Results for Propranonol

Thirty-one effect sizes were calculated on a total of 30 studies for Propranonol treatment. One study, Lucking et al. (1988), reported data on two separate trials so two effect sizes were calculated from this study. Studies ranged in publication years from 1972 to 2000. The meta-analysis on Propranonol included 3,247 total observations, with 1,987 observations in the control group and 1,260 observation in the treatment group. Twenty-three (74.2%) of the studies compared Propranonol to a placebo group. The other eight used a control group rather than a placebo group. Reduction in the frequency of headache was the most consistently reported outcome measure (27 of 31). Doses ranged from 60-240 mg a day, with the mode being 180 mg a day. Modal active treatment length was 12 weeks (range 4 to 52 weeks). Mean Jadad quality score was 3.22, mean study quality score was 2.48, mean mortality rate was 22%, and the study population was 74% female (mean).

Ten of the studies were parallel designs, the other 21 were cross-over designs. Three articles provided enough information to calculate the correlation between groups for cross-over calculations. These were Borgensen et al. (1974) =.82; Mikkelsen et al. (1986) = .81; and Stensrud and Sjaastad (1976) = .74. The lowest correlation (.74) was substituted into the 18 other cross-over studies to account for the cross-over design issue.

Twenty-five studies provided enough continuous data to calculate standardized mean differences. Odds ratios were calculated on the remaining six studies. Summary statistics are listed below in Table 5 by category.

Regression analysis indicated that unbiased effect sizes varied systematically by study design type, with cross-over studies producing significantly higher effect sizes than

Table 5

Findings for Propranolol

Statistic	Continuous data	Binary data
Q statistic	54.2989 $p = 0.0004$	16.4264 $p = 0.0057$
REM summary effect size REM summary CI (95%)	SMD = .68 (.54 - .81)	ln(OR) = 1.58 (1.08 - 2.07)
Variance	0.0044	0.643
Standard deviation	0.06667	0.2536
Rosenthal's file drawer	1122 Credible threat if < 135 $p = 0.0000$ $z = 13.2711$	146 Credible threat if < 40 $p = 0.0000$ $z = 9.85$

Note. SMD = Overall standardized mean difference score; ln(OR) = Overall natural logarithm of the odds ratio.

parallel design studies (R^2= .238, F = 7.89, p =.011). A significant correlation was noted between unbiased effect size and study design type (r = .514), indicating that there is a significant correlation between study design (cross-over or parallel) and effect size. See Appendix G for full reporting of regression analyses and partial correlations.

Results for Flunarizine

Eighteen effect sizes were calculated on a total of 17 studies for Flunarizine treatment. One study, Lucking et al. (1988), reported on two separate trials so two effect sizes were calculated from this study. Studies ranged in publication years from 1982 to 2002. The meta-analysis on Flunarizine included 1,702 total observations. With 1,030 observations in the control group and 672 observation in the treatment group. Eight (44.4%) of the studies compared Flunarizine to a placebo group. The other 10 used a control group or baseline rather than a placebo group. Frequency measures were the most consistently reported measure (15 of 18). Doses ranged from 5-10 mg a day with the mode being 5 mg a day. Modal active treatment length was 14 weeks (range was 4-24 weeks). Mean Jadad quality scores was 2.83, mean study quality score was 2.50, mean mortality rate was 13% and the study population was 75% female (mean).

Fifteen of the studies were parallel designs, the other two were cross-over designs. Fifteen studies provided enough continuous data to calculate standardized mean differences. Odds ratios were calculated on the remaining three studies. Summary statistics are listed below in Table 6 by category.

Table 6

Findings for Flunarizine

Statistic	Continuous data	Binary data
Q statistic	30.9889 $p = 0.0056$	1.6110 $p = 0.4469$
REM summary effect size	SMD = .68	ln(OR) = 1.4942
REM summary CI Interval (95%)	(.53-.83)	(0.73-2.25)
Variance	0.0057	0.152
Standard deviation	0.075	0.39
Rosenthal's file drawer	496 Credible threat if < 85 $p = 0.0000$ $z = 11.4606$	8 Credible threat if < 25 $p = 0.0001$ $z = 3.7247$

Note. SMD = Overall standardized mean difference score; ln(OR) = Overall natural logarithm of the odds ratio

Regression analysis indicated that unbiased effect sizes varied systematically by treatment length, with treatments that were longer producing significantly higher effect sizes ($R^2 = .587$; $F = 17.07$, $p = .001$). A significant correlation was noted between unbiased effect size and treatment length ($r = .733$). See Appendix H for full reporting of regression analyses and partial correlations.

Results for Divalproex Sodium

Seven effect sizes were calculated on Divalproex Sodium treatment. Studies ranged in publication years from 1991 to 2001. The meta-analysis on Divalproex Sodium included 553 total observations, with 213 observations in the control group and 340 observation in the treatment group. Six (85.7%) of the studies compared Divalproex Sodium to a

placebo group. The remaining study used a control group rather than a placebo group. All studies reported frequency outcome measures. Doses ranged from 500-2000 mg a day. Doses were typically individually adjusted to be within optimal serum levels (50-100mg/ml). Modal active treatment length was 12 weeks (range 6 to 24 weeks). Mean Jadad quality scores was 2.57, mean study quality score was 2.14, mean mortality rate was 16.2%, and the study population mean was 83.7% female.

Five studies were of a parallel design and the other two studies were cross-over designs. None of the articles provided enough information to calculate the correlation between groups for cross-over calculations. Therefore, one cross-over study that provided continuous data (Hering & Kuritzky, 1992) was discarded.

One parallel design study provided enough continuous data to calculate standardized mean differences (Kinze et al., 2001). The standardized mean difference was 1.66. A meta-analysis was not performed on this single effect size. Odds ratios were calculated on the remaining six studies. Summary statistics are listed below in Table 7.

Regression analysis indicated that unbiased effect sizes did not systematically vary by study design type, treatment length, quality scores, or by type of control groups used ($R^2 = 1.00$). No significant correlations were noted between any variable and unbiased effect size. See Appendix I for full reporting of regression analysis and partial correlations.

Results for Thermal Biofeedback

Eighteen effect sizes were calculated on 15 studies for thermal biofeedback.

Table 7

Findings for Divalproex Sodium

Statistic	Binary data
Q statistic	4.5794 $p = 0.4693$
REM summary effect size REM summary CI (95%)	ln(OR) =1.65 (1.20-2.11)
Variance	0.0535
Standard deviation	0.2314
Rosenthal's file drawer	80 Credible threat if < 40 $p = 0.0000$ $z = 7.40$

Note. ln(OR) = Overall natural logarithm of the odds ratio.

Three studies (Blanchard et al., 1982, 1991; Jurish et al., 1983), were multi-arm studies that allowed for two calculations to be made from each study. The thermal biofeedback area of treatment can be broken into two protocols of treatment, one that includes explicit relaxation training and one that does not. In this sample 14 effect sizes were calculated from studies that explicitly taught some version of relaxation training (e.g., progressive muscle relaxation, autogenic training) and 6 effect sizes were calculated from studies that did not. Analysis of the findings indicate that results did not significantly vary with the addition of relaxation to the standard thermal biofeedback protocol. Thus, results are aggregated across these two categories.

Studies ranged in publication years from 1978 to 1997. The meta-analysis on thermal biofeedback included 576 total observations, with 261 observations in the control

group, and 315 observations in the treatment group. Fourteen (70%) of the studies compared thermal biofeedback to another treatment. The other four studies compared the treatment to a wait list group rather than an alternate treatment group. Ten effect sizes were calculated from studies that compared thermal biofeedback to a control group. The other eight effect sizes were calculated from within-group comparisons based on studies that compared multiple nonpharmacological treatments. Effect sizes were calculated separately for within-group studies and between group studies to identify if effects were significantly different by study design. Studies that were compared to control groups had an average effect size of .66 and studies that were comparison treatment designs with pre-post measures averaged an effect size of .64. This suggests that comparison treatment studies and controlled outcome studies for migraine prophylaxis are measuring a similar effect. Thus, these two groups were summed for summary effect size calculation. Headache Index measures were the most consistently reported measure (14 of 18). Number of sessions ranged from 2 to 22 with the mode being 12. Modal treatment length was 6 weeks (range 5 to 32 weeks). Mean Jadad quality scores was 1.66, mean study quality score was 2.66, mean mortality rate was 15.22%, and the study population was 81% female (mean).

All eighteen effect sizes were derived from parallel designs and fourteen studies provided enough continuous data to calculate standardized mean differences. Odds ratios were calculated on the remaining four studies. Summary statistics are listed below in Table 8 by category.

Table 8

Findings for Thermal Biofeedback

Statistic	Continuous data	Binary data
Q statistic	2.8813 $p = 0.9983$	3.1559 $p = 0.3682$
REM summary effect size REM summary CI (95%)	SMD = .60 (.40-.79)	ln(OR) = 1.24 (1.08-2.07)
Variance	0.0098	0.148
Standard deviation	0.09899	0.03856
Rosenthal's file drawer	121 Credible threat if < 80 $p = 0.0000$ $z = 6.0671$	8 Credible threat if < 30 $p = 0.0004$ $z = 3.3416$

Note. SMD = Overall standardized mean difference score; ln(OR) = Overall natural logarithm of the odds ratio.

Regression analysis indicated that unbiased effect sizes did not vary systematically ($R^2 = .928$, $F = 3876$, $p = .146$). Variance for both binary and continuous data is very low indicating that effect sizes were fairly uniform and did not vary by study characteristics. No significant correlations were noted between unbiased effect size and any study variable. See Appendix J for regression analyses and partial correlation tables.

Results for Relaxation Therapy

Five effect sizes were calculated for relaxation therapy. Treatments for relaxation therapy included progressive muscle relaxation, autogenic training, and meditation. Studies ranged in publication years from 1978 to 1990 (none after 1990 qualified for inclusion in the meta-analysis). The meta-analysis on relaxation therapy included 186

total observations, with 99 observations in the control group and 87 observations in the treatment group. All five studies compared relaxation therapy to another treatment. Four effect sizes were calculated from studies that compared thermal biofeedback to a control group, while the other effect size was calculated from with-in group comparisons. Headache Index measures were the most consistently reported measure (4 of 5). Number of sessions ranged from 8 to 16, with the mode being 9. Modal treatment length was eight weeks (range 5-37 weeks). Mean Jadad quality score was 1.83, mean study quality score was 2.81, mean mortality rate was 13.8%, and the study population was 81% female (mean).

All effect sizes were derived from parallel designs. Four of five studies provided enough continuous data to calculate standardized mean differences. An OR was calculated on the remaining study. The OR for Daly et al. (1983) was 4.5584 with a confidence interval of (0.88 to 23.37). The Log Odd was 1.51. Continuous data summary statistics are listed in Table 9 below. Partial correlations for relaxations treatments can be found in Appendix K.

Due to the small number or studies included in the meta-analysis a regression analysis and partial correlations were not conducted.

Results for Combination Therapies

The results for the combination therapies include two major treatment categories: EMG biofeedback and thermal biofeedback plus relaxation plus cognitive therapy (TBF + R + Cog). Results will be presented on EMG biofeedback and then on TBF + R + Cog.

Table 9

Findings for Relaxation Therapies

Statistic	Continuous data
Q statistic	6.1586 $p = 0.9770$
REM summary effect size REM summary CI (95%)	SMD = .75 (.37 - 1.13)
Variance	0.0375
Standard deviation	0.1937
Rosenthal's file drawer	17 Credible threat if < 30 $p = 0.0000$ $z = 4.4026$

Note. SMD = Overall standardized mean difference score.

Results for EMG Biofeedback Combination Therapies

The selection criteria for EMG biofeedback resulted in a low number of studies being kept for inclusion. Five different studies were identified for inclusion. Studies that met the inclusion criteria for EMG biofeedback ranged in publication years from 1971 to 1994. The meta-analysis on EMG biofeedback included 97 total observations, with 51 observations in the control group and 46 observations in the treatment group. Three (60%) of the studies compared EMG biofeedback to another treatment. The other two studies compared the treatment to a wait list group rather than to an alternate treatment group. Four effect sizes were calculated from studies that compared EMG biofeedback to a control group, while the other effect size was calculated from with-in group comparisons. Headache Index measures were reported in three of five studies whereas

frequency measures were reported in two studies. Number of sessions ranged from 8 to 16 with the mode being 12. Modal treatment length was 5 weeks (range 5-16 weeks). Mean Jadad quality scores was 1.45, mean study quality score was 2.62, mean mortality rate was 15.75%, and the study population was 88% female (mean). All effect sizes were derived from parallel designs.

A total of five effect sizes were calculated for EMG biofeedback. EMG biofeedback comes in two predominate forms; (a) EMG + Relax, or (b) treatment combined with thermal biofeedback and relaxation training (EMG + TBF + R). In this sample three effect sizes were calculated on EMG + TBF + R, and two effect sizes were calculated from studies on EMG + Relax. Analysis of these five studies showed that the two different protocols have distinctly different effect sizes and thus the results are reported separately. Odds ratios were calculated on all studies. The results shown in Table 10 are for the only two controlled studies available for EMG + Relax. Table 11 presents the findings for EMG + TBF + R.

The following results are on the three controlled studies that reported treatment data on EMG + TBF + R.

Regression analysis was not performed on these results due to the low study numbers. However, effect sizes were fairly uniform within treatment protocols indicating that effect sizes were unlikely to vary by study characteristic.

Results for Thermal Biofeedback plus Relaxation Therapy plus Cognitive Therapy

A total of four different studies were identified for inclusion. Studies that met the

Table 10

Findings for EMG + Relax

Statistic	Binary data
Q statistic	0.8119 $p = 0.3675$
REM summary effect size REM summary CI (95%)	ln(OR) = 2.47 (4.0-.93)
Variance	0.6184
Standard deviation	0.786
Rosenthal's file drawer	4 Credible theat if < 20 $p = 0.0005$ $z = 3.2693$

Note. ln(OR) = Overall natural logarithm of the odds ratio.

Table 11

Findings for EMG + TBF + R

Statistic	Binary data
Q statistic	0.0086 $p = 0.9957$
REM summary effect size REM summary CI (95%)	ln(OR) = 1.1064 (2.13-.075)
Variance	0.2765
Standard deviation	0.526
Rosenthal's file drawer	1 Credible threat if < 20 $p = 0.0190$ $z = 2.0759$

Note. ln(OR) = Overall natural logarithm of the odds ratio.

inclusion criteria for TBF + R + Cog ranged in publication years from 1979 to 1990. The meta-analysis on TBF + R + Cot included 112 total observations, with 49 observations in the control group and 63 observations in the treatment group. All of the studies compared TBF + R + Cog to another treatment. Three effect sizes were calculated from studies that compared TBF + R + Cog to a control group, while the other effect size was calculated from with-in group comparisons. Headache Index measures were reported in two of four studies, frequency measures in two of four studies. Number of sessions ranged from 5 to 18, with the mode being 5. Modal treatment length was five weeks (range 5 to 12 weeks). Mean Jadad quality score was 1.75, mean study quality score was 2.50, mean mortality rate was 14.25%, and the study population was 76% female (mean).

All effect sizes were derived from parallel designs. Three of four studies provided enough continuous data to calculate standardized mean differences. An OR was calculated on the remaining study. The OR for Lake et al. (1979) was 2.0370 with a confidence interval of (0.18-21.94). The Log (OR) was 0.71. Continuous data summary statistics are listed in Table 12 below.

Due to the small number of studies included in the meta-analysis, a regression analysis and correlation analysis was not conducted.

Results of Comparisons Between Summary Statistics

Results for comparisons between meta-analyses will be presented in two ways. First, results are presented in table format for visual comparison. Deeks (2002) suggested that if

Table 12

Findings for TBF + R + Cog

Statistic	Continuous data
Q statistic	1.1431 $p = 0.5647$
REM summary effect size REM summary CI (95%)	SMD = 0.7260 (1.13-.31)
Variance	0.0442
Standard deviation	0.2102
Rosenthal's file drawer	6 Credible threat if < 25 $p = 0.0005$ $z = 3.3080$

Note. SMD = Overall standardized mean difference score.

summary statistics fall within each others' confidence intervals then treatments are roughly equal. Thus, the effect sizes for treatments with more than one continuous outcome are presented in Table 13. Then, results of an ANOVA (Table 14) conducted on the summary statistics are reported. The summary results for ln(OR) data collected on treatments with more than one binary outcome are then reported in Table 15, with ANOVA results in Table 16.

Excluded in the summary analysis is the single continuous effect size found for Divalproex Sodium ($d = 1.66$) and the single OR found for relaxation therapies (ln (OR) = 1.51) and TBF + R + Cog (ln (OR) = 0.71). Because these are single effect sizes they do not represent a body of literature and, thus, are not included.

It should be noted that due to the nature of the two different types of data

Table 13

Summary Statistics for Continuous Outcomes

Treatment	N	REM summary effect size (SDM)	95% confidence interval
Propranolol	25	0.68	.54- .81
Flunarizine	15	0.68	.53- .83
Thermal biofeedback	14	0.60	.40- .79
Relaxation therapy	4	0.75	.37-1.13
Combination therapy TBF + R + Cog	3	0.72	.31-1.13

Table 14

ANOVA for Continuous Summary Statistics

	Sum of squares	*df*	variance	*F*	*p* value
Between-group	0.105	4	0.03	2.9024	0.0297
Within-group	0.51	56	0.01		
Total	0.61	60			

(continuous and binary) not all outcomes could be directly compared to each another. Only one type of effect size was calculated for each study. If the researchers provided binary data, an OR was calculated; if continuous data were provided, a standardized mean difference effect size was calculated. Therefore, in some cases a treatment category (e.g., Propranolol) will have some OR effect sizes and some standardized mean difference effect

Table 15

Summary Statistics for Binary Outcomes

Treatment	*N*	REM summary effect size ln(OR)	95% confidence interval
Propranonol	6	1.58	1.08-2.07
Flunarizine	3	1.49	.73 -2.25
Divalproex Sodium	5	1.65	1.20-2.11
Thermal biofeedback	64	1.24	1.08-2.07
Combination therapy			
EMG + Relax	2	2.47	.93-4.0
EMG+ TBF + Relax	3	1.11	.075-2.13

Table 16

ANOVA for Binary Summary Statistics

Source	Sum of squares	*df*	variance	*F*	*p* value
Between-group	2.29	5	0.46	3.1775	0.0314
Within-group	2.60	18	0.14		
Total	4.89	23			

sizes. Both types of data were found for Propranolol, Flunarizine, and thermal biofeedback. Binary data only were found for Divalproex Sodium, EMG biofeedback plus Relaxation, and EMG biofeedback plus thermal biofeedback plus relaxation therapy. Only continuous data were used to calculate effect sizes for relaxation therapy and thermal

biofeedback plus relaxation plus cognitive therapy. The result of including two types of data is that the outcomes for some treatments types are not directly compared to other treatment types. For example, Divalproex Sodium has one continuous effect size (mentioned above) that was excluded and six binary effect sizes. The six binary effect sizes are summed and reported in Table 15, resulting in Divalproex Sodium only being compared to the other treatments that have binary outcomes. This excludes relaxation therapy, which reported five continuous outcomes and is summed in Table 13. Thus, Divalproex Sodium and relaxation therapy are not directly compared.

Indirect comparisons can be made by using the treatments that provided both types of data. For example, Divalproex Sodium (binary) and relaxation therapy (continuous) can be indirectly compared by directly comparing each to Propranolol (which provides both binary and continuous outcomes). Through visual inspection it can be determined that Propranolol and Divalproex Sodium have similar effect sizes and that Propranolol and relaxation therapy have similar effect sizes; thus, Divalproex Sodium and relaxation therapy are likely to have the similar effect sizes.

The results from the one-way ANOVA to compare continuous summary effect sizes revealed a significant effect $F = 2.902$, $p = .0297$ (see Table 14). Post hoc Bonferroni adjusted comparisons revealed no significant individual group comparisons (see Table L1 in Appendix L).

The results from the one- way ANOVA to compare groups revealed a significant effect $F = 3.1775$, $p = .0314$ (see Table 16). Post hoc Bonferroni adjusted comparisons revealed no significant individual group comparisons (see Table L2 in Appendix L).

CHAPTER VI

DISCUSSION

The current meta-analysis was conducted to answer four questions: (a) How do the effectiveness rates of treatments compare to each other and between groups (pharmacological and nonpharmacological)? (b)What sample characteristics are significantly correlated with outcomes? (c) Do outcomes vary if broken into two levels of empirical evidence? (d) What information is available regarding short and long term effectiveness rates for pharmacological and nonpharmacological studies? Each of these questions will be addressed in turn. A discussion of the practical uses of these findings follows. Then, limitations of the current findings will be discussed. Finally, suggestions for future research in the area of migraine headache are proposed.

How the Effectiveness Rates of Treatments Compare to Each Other and Between Groups

The results of the present meta-analysis suggest that five of the six evaluated treatments have very similar effect sizes. In addition, the findings suggest that three of the summary effect sizes (EMG biofeedback, relaxation therapy, and thermal biofeedback plus cognitive therapy) are tentative due to a serious threats to validity from the file drawer phenomenon. The evidence for these conclusions are reviewed below in the following order (a) visual comparison of effect sizes, (b) Rosenthal's file drawer findings, (c) inferential statistical findings, (d) findings for EMG biofeedback plus relaxation, (e)

findings for EMG biofeedback plus thermal biofeedback plus relaxation, and (f) the summary of the outcomes and possible explanations for findings.

Comparisons by Visual Analysis

The current meta-analysis compiled treatment effectiveness data on six prophylactic treatments for migraine. The summary statistics presented represent the results for the reduction of frequency or headache index for each respective treatment of migraine headache. Summary statistics for continuous data range from .60 to .75. The lowest summary effect size was found for thermal biofeedback treatment and the highest summary effect size was found for relaxation therapy. All other treatments (Propranonol, Flunarizine, Divalproex Sodium, and mixed treatments) had summary effect sizes that fell within this range. All summary statistics fall within each other's confidence interval. According to Gall, Borg, and Gall (1996), these effect sizes are in the large range of treatment effectiveness.

Evaluation of the binary summary statistics also show two patterns. The majority of the log (OR) statistics fall in a narrow range of effect sizes (1.49-1.65). Again, those effect sizes that were clustered together all fell within one another's confidence intervals. There are two exceptions for this pattern and they both involve EMG biofeedback. The EMG biofeedback treatment has two protocols, one that includes thermal biofeedback and one that does not. Analysis of the results revealed that when the protocols were split, the EMG biofeedback treatment that included thermal biofeedback was low, log (OR) = 1.10, and the EMG biofeedback that did not include thermal biofeedback was substantially

higher, log (OR)= 2.47. The findings for EMG biofeedback represent the only effect size that appears to be significantly different from the other treatments analyzed. Possible reasons for all of these findings will be presented below following the discussion of the inferential statistical results.

Rosenthal's File Drawer

Rosenthal's file drawer ratings for each treatment indicates that the file drawer threat is a valid threat to the following treatments, Flunarizine (binary data only), thermal biofeedback (binary data only), relaxation treatments, EMG biofeedback plus relaxation, EMG biofeedback plus thermal biofeedback plus relaxation, and thermal biofeedback plus relaxation plus cognitive therapy. The first two treatments, Flunarizine and thermal biofeedback, are not seriously threatened by the file drawer phenomenon because the continuous data provides more than enough findings to support the conclusions in the published research. Therefore, the results for Propranonol, Flunarizine, Divalproex Sodium, and thermal biofeedback are considered to be stable results that could not be easily threatened by nonpublished studies that show no difference between the treatments and control groups.

However, the other treatment results (EMG treatments, relaxation treatments, and thermal biofeedback plus relaxation plus cognitive therapy) are potentially threatened by nonpublished studies that support the null hypothesis. Thus, less confidence can be placed in the findings for EMG treatments, relaxation treatments, and thermal biofeedback plus relaxation plus cognitive therapy. This is largely due to the small number of studies used

to find the summary statistic in each of these treatments. Of particular concern are the results for the EMG biofeedback plus thermal biofeedback plus relaxation treatment that indicates that only one unpublished article (that finds support for the null hypothesis) could be a valid threat to these results. So, while the present findings are suggestive of the basic trend for these treatments, there is not enough data to form solid conclusions about their treatment effect without further replication of outcome findings.

Comparisons by Inferential Statistics

Inferential statistics were used to compare effect sizes to one another. ANOVA procedures were conducted to attempt to identify statistically significant differences between the summary statistics. Results from the ANOVA indicated that there was a significant difference between the summary statistics for both the continuous data ($F = 2.90$; $p = .029$) and binary data ($F = 3.17$; $p = .031$). Post-hoc analysis with Bonferroni corrections were performed. None of the post-hoc analyses reached significance. This is likely due to the large number of comparisons that were conducted between the groups. By performing approximately 10 comparisons for each group of outcomes the significance value with Bonferroni correction becomes $p < .005$ (continuous) or $p < 003$. (binary). This is an extremely stringent test and these treatments did not reach the criterion.

It should be noted that several authors warn that using inferential statistics is highly inappropriate for meta-analyses (e.g., Abrami, Cohen, & Apollonia, 1988; Schmidt, 1992). They argue that the nature of effect sizes violate the assumptions of inferential statistics (i.e., they are not samples of the population, they are the population; they do not meet the

assumptions of normal distributions, etc.). In addition, they report that inferential statistics are inappropriate for measurement based on the individual unit; effect sizes that are weighted (as are all summary statistics used under the random effects model) result in a statistic that is based on the individual. The current meta-analysis is likely to have been effected by these issues. In the present meta-analysis a specific population of interest was defined a priori (i.e., through the use of inclusionary rules such as migraine prophylactic studies that examined one of six treatments, providing enough data to calculate an effect size, etc.). A systematic search strategy was used to attempt to obtain all studies that made up the defined population. As a result, the use of inferential statistics would be inappropriate in the current meta-analysis to the same extent that the efforts to obtain *all* articles were successful. Thus, the results of the ANOVA's and post-hoc Bonferonni calculations in the current meta-analysis may be invalid due to assumption violations and should be interpreted with caution.

EMG Biofeedback Plus Thermal Biofeedback

The low ln (OR) summary statistic for EMG biofeedback plus thermal biofeedback is not as interpretively important as the meaning of this treatment's confidence interval. When a ln (OR) confidence interval encompasses zero it can represent that a treatment has no greater effect than the comparative treatment (in this case, a control group). It should be noted that the EMG biofeedback plus thermal biofeedback confidence interval is the only treatment that has a confidence interval that approaches zero (.075). If the true treatment effect for EMG biofeedback plus thermal biofeedback lies in the lower end of

the confidence interval, it would indicate that the odds of an individual having a successful outcome with this treatment may be equal to the odds of a placebo outcome.

Statistically, the wide confidence interval found for EMG biofeedback plus thermal biofeedback (.075-2.14) indicates that there is more variability in this summary effect size than in some of the others. Close examination of the findings for EMG biofeedback plus thermal biofeedback indicates that the three primary effect sizes that make up the overall summary effect size are similar to each other and to the summary effect size (ln (OR) = 1.10, 1.06, and 1.16; REM summary effect size ln (OR) = 1.11). Thus, the variation between studies is quite low. However, the confidence interval calculated using the random effects model takes into account within study variation as well as between study variation. Thus, this large confidence interval was likely effected by the within study variance. Of particular note is the within study variance found for Largen et al. (1981), the standard deviation of within study variance for this study was 1.21, the two other effect sizes for EMG biofeedback plus thermal biofeedback had standard deviations at .80 and .85. Indicating that the Largen study introduced much of the within study variability found in this summary effect size.

The high level of variability in Largen et al. (1981) may be explained by the sample sizes or the small effects that were reported. The Largen study has particularly small sample sizes (experimental group $N = 6$, control group $N = 5$); small sample sizes can introduce more variability into the ln (OR) effect size calculations than larger sample sizes. In addition, the Largen study had only three successes in the experimental group (three of six), indicating that the odds of success in this condition were not greatly in favor

of EMG biofeedback plus thermal biofeedback. These two added sources of variation may have widened the ln(OR) effect size confidence interval and may explain this finding in the current meta-analysis.

The summary effect size in for EMG biofeedback plus thermal biofeedback (ln (OR) = 1.11) is also low. The poor outcome for this type of treatment may have been influenced by several major factors; number of studies used to calculate the treatment effect, study characteristic, or a lack of a true treatment effect. Each of these possibilities is explored below.

The summary statistic for EMG biofeedback plus thermal biofeedback is based on a small number of studies ($N = 3$). However, EMG biofeedback plus relaxation is based on only two studies, and the resulting summary statistic does not display a problematic confidence interval. Second, the poor outcome for EMG biofeedback plus thermal biofeedback may represent a study design flaw that negatively impacts the treatment outcome. An analysis of the studies used to find the summary statistic for EMG biofeedback plus thermal biofeedback shows that all three studies used control groups instead of wait list groups. It is possible, given difficulty finding appropriate attention placebo groups, that these control groups were flawed and systematically bias the outcome of this analysis.

Examination of the individual studies indicates that one study used thermal biofeedback with temperature cooling as a control and the other two used groups that were told to relax daily. Both could bias findings by providing a treatment effect in the control groups. Temperature cooling has a mixed history of effectiveness. Several authors

(Blanchard et al., 1997; Gauthier, Bois, Allaire, & Drolet, 1981) have reported that temperature cooling biofeedback can be equally effective to thermal warming biofeedback. However, it is a commonly held belief that thermal biofeedback with a focus on temperature cooling is counter intuitive to migraine pathophysiology and is ineffective (Largen et al., 1981). The other two primary studies used to establish the effect size for EMG biofeedback plus thermal biofeedback had control groups that were "self-relax" groups. The self-relax groups were instructed to "relax" 10-15 minutes a day. There is a known treatment effect for structured daily relaxation technique. While these studies did not use structured techniques, the self relaxation instructions given to subjects may have resulted in a small treatment effect for the control group. Given that all three of the control groups have at one time been considered an active treatment on migraine headache, the "control" group participants in these studies may have benefitted to some degree from the instructions they were given, thereby reducing the difference between the two groups and decreasing the resulting OR. Anecdotal review of the studies used to calculate the summary effect size for EMG biofeedback plus thermal biofeedback lends support to this hypothesis. Each study in the EMG biofeedback plus thermal biofeedback treatment group reported that 20-30% of participants in control groups reached the success criteria (50% reduction in frequency of headache). This can be compared to the studies used in the EMG biofeedback plus relaxation treatment group who each report between 16-18% of participants in the control groups met the success criteria.

Finally, the low effect size found for EMG biofeedback plus thermal biofeedback may reflect a true treatment effect. The three studies that utilized EMG biofeedback plus

thermal biofeedback all included initial sessions of EMG biofeedback that were followed by thermal biofeedback training. It is possible that the initial EMG biofeedback training in some way interfered with the typically good outcome of the thermal biofeedback treatment.

EMG Biofeedback plus Relaxation Techniques

The ln (OR) summary statistic for EMG plus relaxation techniques is 2.47. This is the highest binary summary statistic found in the current meta-analysis. This summary statistic also has a large confidence interval (.93 - 4.0) The summary statistic is based on only two studies, they both resulted in high ORs but are somewhat different from each other (primary study ln (OR) = 2.11 and 3.86), which would account for the confidence interval. The high level of treatment effect noted for EMG biofeedback plus relaxation training may have been influenced by the fact that both studies have relatively early publication dates (1979 and 1983), both included mixed headache patients, and both had very low quality ratings. In addition, one of the studies (Daly et al., 1983) was not randomized, the subjects were assigned to groups based on the severity of their symptoms. These study characteristics may have introduced bias that inflated the EMG biofeedback plus relaxation treatment effect size.

Summary of Outcomes

The findings of this meta-analysis indicate that five of the treatments evaluated produce moderate to large effect sizes. The one exception to this finding is the relatively

large effect size for EMG biofeedback plus thermal biofeedback. However, it is likely that the findings for EMG biofeedback were unduly influenced by study characteristics. This study's findings regarding effect sizes for these five treatments for migraine are nearly identical to those found by the Duke University study conducted for the Department of Commerce (Duke University and Center for Clinical Health Policy Research, 1999a, 1999b). These studies reflect the same findings although some of the statistical and selection procedures varied. This "replication" of the Duke University study offers strong support for the finding that these treatments offer similar moderate-to-large rates of effectiveness.

It is curious to find that five different treatments based on highly varied theories and approaches have approximately the same treatment effect for migraine prophylaxis. Results based on OR indicate that between 40-70% of the subjects in clinical trials reach clinical reduction of symptoms (i.e., 50% reduction in frequency or headache index) no matter what treatment they receive. These findings beg the question, "Why do these treatments all have approximately the same effect when they are so different?" While there is no succinct answer to this question, there are important factors that may contribute to this finding. The first factor impacting the treatment effects may be compliance. Second, migraine, as currently defined, may represent a set of disorders, some of which are amenable to the types of treatments currently being used and others that are not being affected.

Compliance

Compliance has long been an issue for all long term treatments. Mullerners, Whitmarsh, and Steiner (1998) indicated that compliance in migraine prophylaxis is a major issue. They report that through the use of a computer monitoring system (secretly inserted into the lids of patient pill bottles) they were able to identify the actual compliance of subjects receiving migraine prophylactic treatment. Mulleners et al. (1998) reported that the used-on-schedule compliance rate (even in a controlled study) was, on average, 30% for participants who were to take medication two or three times a day ($N = 18$). Participants who took medications once a day averaged 66% used on schedule compliance rates ($N = 11$). The authors suggested that these low compliance rates significantly altered the effect of medical treatments intended to reduce the frequency of migraine. It is theoretically possible to extend these results to the nonpharmacological treatments for migraine. If patients have difficulty accurately taking a medication one to three times a day, it is questionable if they will adequately practice biofeedback techniques or relaxation techniques at 10-15 minutes per day.

Mulleners et al. (1998) also found that patients were remarkably inaccurate about reporting to researchers how many pills they took. Even with a known compliance check (counting remaining pills) they found that there was a large difference between the amount of pills actually taken and the amount of pills reported to be taken. This finding seriously calls into question the results of studies reporting nearly perfect compliance with treatment regimes.

If compliance rates are as low as suggested by Mulleners et al. (1998), then it may

account for why none of these varied and sometimes intensive treatments have reached a desired level of effect or been able to differentiate one treatment from another. In addition, these findings have "real world" implications for how usable daily treatments are for the general population. If subjects cannot partake of treatment in a prescribed manner during a highly controlled and structured situation, it is not very likely that they will appropriately engage in treatment under everyday living situations.

Migraines as a Broad Category

The current meta-analysis finds that treatments ranging from structured guided relaxation to calcium channel blockers have similar effects in migraine prophylaxis. It is possible that the finding of equivalent treatments effects indicates that more than one disorder is being treated by the migraine prophylactic treatments. Thus, chronic migraine headaches may be indicative of several different physiological and psychological pathways. If different disorders are currently grouped into the "migraine" category, and each treatment addresses a different underlying condition, it could account for the findings of similar effect sizes.

Recent genetic research may offer some support of this hypothesis. Although identifying genetic markers for diseases is still in its infancy as a science, some interesting findings regarding migraine headaches have already been reported. For example, Goadsby (1997) reported that familial migraine has been linked (in some families) to chromosome 19p13. Goadsby went on to report that preliminary findings suggest that different families may be linked to other chromosomal regions. Of particular interest has been the finding

that for those who have a link to chromosome 19p13 some also have links to lower channel genes that are voltage-gated ionophores for Na+ and or K+. Flunarizine and Divalproex Sodium both have Na+ and Ca+ ion effects. Thus, these treatment may only work for those who have the particular genetic profile that is associated with Na+ deficits. As the area of genetic research broadens it may identify other familial deficits that best account for treatment effects of these and other types of treatments.

While different types of migraines may exist, it appears that migraines are likely to all share common physiological pathways. If they did not, the current treatments would be unlikely to treat the noted large numbers. Goatsby (1997) reviewed experimental human and animal research and identified the common pathways that pharmacological agents share. Most prominent were the 5HT receptors and the calcium or sodium channels. Goatsby reported that most prophylactic pharmacological agents impacted at least one of the seven subclasses of 5HT receptors. Goatsby reported that Methysergide, Pizotifen, Propranolol, Amitryptiyline, Imipramine, and Flunarizine had documented effects on the 5HT system. Alternatively, Goatsby noted that Flunarizine and Divalproex Sodium had documented effects on the active sodium and calcium ion channels.

Nonpharmacological treatments have less experimental data available in the research. However, several authors hypothesize that nonpharmacological treatments promote biological homeostasis that reduces the likelihood that the above mentioned systems will be negatively impacted (e.g., Blanchard et al., 1978; Feurstein, Bortolussi, Houle, & Labbe, 1983). Holroyd (2002) reported that cortical excitability has been strongly linked to the onset of migraine. This excitability (related to neurogenic

inflamation that sensitizes nerve endings) can be altered through psychophysiological interventions that seek to quiet the system and thereby reduce brain stem activity (Holroyd). The brain stem has been implicated as the "migraine generator" due to its role in pain reception and vascular control (Weillner et al., 1995). Holroyd suggests that nonpharmacological treatments may be nonspecifically reducing the reactivity of the brain stem, making this a likely choice for the common pathway mechanism that the nonpharmacological treatments share.

Goatsby (1997) focused mostly on pharmacological treatments; however, his comments on available research can be applied to both areas (pharmacological and nonpharmacological treatments).

> The understanding of the action of preventative drugs is at a relatively immature stage. Recent developments in pharmacology and studies of pathophysiology of migraine have provided a substrate around which concepts can be developed. The actions and indeed the locus of action of the preventatives is crucial since these drugs are likely to point to the basic defect which underlies the process responsible for a migraine attack. (p. 90)

Thus, suggesting that the common pathways that the pharmacological treatments share is a crucial yet undeveloped area of understanding. This statement can be applied to the understanding of the common pathways of nonpharmacological treatments, that are also in their infancy. Future research on how migraine prophylactic treatments work can help develop clearer understanding of migraine pathogenesis and would help to clarify why the treatments in the current meta-analysis have nearly identical effects.

Sample Characteristics That Are Significantly Correlated with Outcomes

The current meta-analysis found that only two treatment groups displayed sample characteristics that significantly correlated with outcome. Study design significantly correlated with outcome for Propranonol and treatment length significantly correlated with outcome for Flunarizine. No other significant findings for study characteristics were found. Notably, study quality was not among those sample characteristics that correlated with outcome. Below, significant findings for Propranonol and Flunarizine are discussed, followed by a discussion of study quality.

Propranonol and Study Design

The regression analysis for Propranonol treatment outcomes for migraine prophylaxis identified that effect size outcome was significantly correlated with study design. The finding indicated that cross-over designs produced higher effect sizes than parallel designs. This finding may be largely due to the method of including cross-over studies in this meta-analysis. The lengthy history of Propranonol as a migraine prophylactic has resulted in years of research examining Propranonol's effects versus other pharmacological agents. This has resulted in a high number of primary articles that have cross-over designs. As previously mentioned, the results from cross-over designs have typically been ignored in meta-analyses. The current meta-analysis attempted to include the results by using a correlation in the calculation of effect size. When a correlation is included in the calculation of effect size the result is always of greater magnitude than

uncorrelated effect size correlations. Thus, it is not surprising to find that the two types of study designs produced systematically different outcomes.

However, researchers have suggested that it is appropriate for cross-over designs to yield more powerful results than parallel designs (Kunkel, 1987). Cross-over designs are considered to hold at least twice the power of parallel designs because participants serve as their own control and their own comparison. The current meta-analysis chose a conservative approach to this issue. Correlations were gathered from all studies providing them and correlations were calculated for all studies that provided individual data. The lowest correlation obtained was substituted for all cross-over studies (even if the original study provided a higher correlation). Thus, the finding that cross-over studies have higher effect sizes may be a product of the statistical procedures used in this analysis. The example of effect sizes for cross-over studies were calculated in Appendix F. It is clear in the example that the effect size is larger when the correlation is added in. Thus, the finding that cross-over designs had higher effect sizes is not surprising given that a priori decision to add more "weight" to these studies by using correlations with all effect size calculations for cross-over designs.

Flunarizine and Treatment Length

The current meta-analysis found that effect sizes increased as length of treatment increased for subjects who were treated with Flunarizine. This is not surprising given that Flunarizine makes plasma levels increase slowly with daily oral administration. Serum levels do not reach a steady state for 4 to 6 weeks (Todd & Benfield, 1989). This indicates

that it takes a longer time to reach a "therapeutic dose" level in the patient. The slow effect of Flunarizine is supported by the current outcome findings. Many of the original research articles on Flunarizine attempted to use the Propranonol research protocol that often treated patients for 8 to 12 weeks. Flunarizine's full effect seems only to begin to appear between 8 and 12 weeks. Thus, later research shifted to a 16- to 20-week protocol, with positive results. This appears to account for the finding that Flunarizine's effect size seem to grow with length of treatment.

Quality of Study

The current study did not identify any statistically significant relationships between quality of study (as measured in this analysis) and outcome. The lack of association between quality of study rating and outcome is likely due to the selection criteria used in this meta-analysis. That is, the selection criteria were fairly rigid and may have selected a fairly homogenous group of articles. There were two quality rating scores obtained for each article included in the meta-analysis. The first was the Jadad Quality Score and the second was an author-developed study quality rating score. Quality of study in the current meta-analysis was defined in two ways: (a) quality of study as defined by Jadad et al. (1996) is defined as how much the authors report the use of high quality study design procedures (i.e., random assignment, double blinding, and description of dropouts), and (b) quality of study is defined by how well the authors addressed concerns related to migraine prophylactic treatment. The scale used to assess the first component of study quality was proposed by Jadad et al. and has been shown to have good inter-rater

reliability and has been empirically validated as an assessment tool for study quality. The authors report that the instrument has been shown to be used consistently by raters regardless of background or training, indicating that it is likely to be valid and reliable in multiple settings including the current one.

The second quality score has a high inter-rater reliablity (reliablility coefficient was found to be .88). However, it has not been empirically validated. The quality score has face validity as it is based on the recommendations of the authors who were reviewed earlier. There it was identified that migraine prophylaxis treatment outcomes were often impacted by the following: type of measurement used (Holroyd & Penzien, 1990), type of control used (Compas et al., 1998), study design (Holroyd et al., 1991), use of run-in/ washout periods (Bogaards & ter Kuile, 1994), multiple setting measurement (Bogaards & ter Kuile), use of a comparison to another active treatment (Chambless & Hollon, 1998), controlling other medication (Onghena & Van Houdenhove, 1992), and measuring compliance or home practice (Holroyd & Penzien).

Each of the above variables were believed to pose threats to validity of reported outcomes in a study on the quality of a migraine prophylaxis outcome. However, the way these variable were rated (as dichotomous, either present or not) and summed (with some variables being seen as necessary and other being seen as important) has not been validated. The two quality of study ratings did reach similar results (no correlation between outcome and study quality) which may indicate that they both are measuring the same broad construct of study quality. Nevertheless, the only conclusion about quality of

study that can be drawn from the current meta-analysis is that quality of study as measured in this meta-analysis did not significantly correlate with treatment outcome.

It should be noted that there were differences between treatments in the mean quality rating score. The following mean quality scores were obtained for each treatment (Jadad Quality Rating Scores are listed first, then study quality rating scores), Propranonol (3.22, 2.48), Flunarizine (2.83, 2.50), Divalproic Sodium (2.57, 2.14), thermal biofeedback (1.66, 2.66), relaxation treatments (1.83, 2.81) EMG (1.45, 2.62) and thermal biofeedback plus cognitive therapy (1.75, 2.50).

The Jadad Quality Score ranges from 0 -5. These scores indicate to what degree the researchers used appropriate research design to rule out threats to validity. The scale rates each article on the following: random assignment (up to 2 points), double-blind design (up to 2 points) and description of drop outs (1 point). Average scores above 3 on this scale are likely to indicate studies of high quality that have relatively few validity threats as a result of study design. The only treatment that exceeded an average Jadad score of 3.0 was Propranonol (3.22). An average score between 2 and 3 represents a group of studies that may have more significant threats to validity as a result of study design or may have had inadequate published reporting of the study design. Both Flunarizine (2.83) and Divalproic Sodium (2.57) fall into this category. A score below 2 on the Jadad Quality Scale indicates a group of articles whose findings may have serious threats to validity due to study design. All the nonpharmacological studies fall into this category, this is due to the lack of "double blinding." None of the behavioral articles received credit for double blinding, reducing the scores of all behavioral studies by at least

2 points. Double blinding is a proposed research technique that is designed to reduce the effect of experimenter/participant expectation on outcome. The ability to actually achieve true double blinding is questionable. It has been argued (e.g., Kirk-Smith & Stretch, 2001; Turner, Jensen, Warms, & Cardenas, 2002) that treatment providers (nurses, doctors, etc.) and participants in pharmacological treatment evaluation studies with double-blind conditions are able to guess which treatment condition they are in (experimental or placebo). Thus, potentially invalidating attempts to double blind the study. Double-blinding becomes even more problematic when nonpharmacological treatments are considered. Some authors (e.g., Lukoff, Edwards, & Miller, 1998) have argued it is nearly impossible to double blind a nonmedical intervention that involves providing an office-based intervention. Whether possible or not the attempts at double blinding in the field of psychological and behavioral research appear to be very low. Sheldrake (1998) reported that in a review of the top psychological journal only 9% of experimental articles use a double-blind methodology. This low use of the double-blind methodology is reflected in the nonpharmacological study of migraine prophylaxis. The nonpharmacological treatment results must be interpreted with caution due to the limited use of double blinding.

The second type of quality rating used in this meta-analysis was developed by the current researcher and was based on previous literature reviews on migraine prophylactic treatments. The Jadad Quality Rating scale focuses on characteristics of standard study design, the second quality rating focused on variables that spedifically pose validity threats in headache outcome studies. For example, did the researchers include a wash-out period in which participants were withdrawn from other medications so that the experimental

treatment alone could account for treatment effect? This quality rating score ranges from 1 to 3, with 3 representing the highest quality score. The quality ratings for the treatments examined were fairly consistent across treatment type. The mean scores ranged from 2.14 to 2.8, indicating that the articles included in this meta-analysis fairly consistently attempted to control for variables that are known to impact treatment outcome in migraine prophylaxis.

Do Outcomes Vary if Broken into Two Levels of Empirical Evidence?

Chambless and Hollon (1998) suggested that the "premiere" type of outcome research compared a treatment that was known to be effective to one that was being tested. They suggested that a treatment reported to be better than placebo revealed nothing about how it compared to other available treatments and thus was not as useful. The current meta-analysis attempted to identify whether this premiere type of research actually provided different information or just more convenient information. To address this issue placebo/control groups were coded separately from comparison group studies. This variable (along with other identified variables of interest) was entered into regression analysis for each treatment to identify if treatment outcome systematically varied by the type of comparison group (control vs. active treatment). The findings of the present meta-analysis indicated that none of the regression analyses included the variable "treatment comparison group" as a significant variable related to outcome (see Appendices G, H, I, and J). Thus, in the area of migraine prophylaxis, the type of outcome information

provided by comparison research studies was similar in magnitude to the information provided in placebo/controlled studies.

In the area of migraine prophylaxis treatment outcome identifying the premiere type of research may depend largely upon how the author reports the results. Migraine prophylactic treatments, as shown in the present meta-analysis, have very similar treatment outcomes. Thus, comparison studies do little to enlighten us about which treatment is better, rather they merely confirm that there is no statistical difference between treatments. If the author chooses to only report nonsignificant results between two active treatments, then the results of comparison research studies cannot be synthesized with other research studies. Such reporting makes comparison research studies less "useable" than placebo/controlled studies. Thus, Chambless and Hollon's (1998) suggestion that comparison studies are the most valuable way to compare studies may only be true when there is comprehensive reporting of results. If this is not the case, articles that provide effect size information are equally useable regardless of whether a comparison group was used or not.

Information Available Regarding Short- and Long-Term Effectiveness

The current meta-analysis coded for length of treatment and length of follow-up. Each treatment varied in the amount of long-term data that was collected. Overall, all six treatments showed that effects of treatment were maintained as long as treatment was maintained. Most treatments examined in the present meta-analysis have studies showing

up to 6 months of treatment with positive outcomes. However, examination of treatment effects after discontinuation of treatment is difficult to determine. The methodological problems of determining long-term effects for a previous treatment for migraine will be discussed below. Followed by a discussion of known long-term effects of treatments evaluated in the present meta-analysis.

Methodological Problems

Data on migraine prophylactic treatments after the treatment has been discontinued are valuable. However, studies that attempt to follow up on a migraine prophylactic treatment face some inherent problems. First, it is unethical to maintain control groups for long periods of time when known treatments exist. Second, without a control group it is impossible to tell if the reported levels of symptomology are related to the previous treatment or the waxing and waning nature of migraine. When long periods of time are involved it is nearly impossible to rule out alternative explanations for reductions of reported symptoms from baseline. Finally, the selection process in follow-up studies is flawed due to unavoidable selection bias. Each of these issues are addressed below.

The issues related to the lack of control groups for long-term outcomes is a significant one. First, it is considered unethical to withhold treatment from a control group once the active phase of the treatment has been discontinued (Sorbi, Tellegen, & DuLong, 1989). Therefore, many researchers offer the active treatment to control subjects after the first phase of the study discontinues. Once the subjects have been treated, there exists no treatment-free group with which to compare the results of patients after a certain amount

of time has elapsed. The absence of a control group precludes the calculation of an effect size for inclusion in a meta-analysis.

Second, the absence of a control group in a long-term prophylactic treatment outcome study for migraines is particularly problematic due to the nature of migraine headaches. Migraine is considered to be a cyclical condition in which cycles of migraines occur and remit in an unpredictable fashion (Couch, 1987). Symptoms of migraines are highly variable and can include intense periods of frequent migraines and extended periods without migraines. Couch reported that between 28-60% of migraine sufferers have symptoms that remit with placebo treatment. Thus, when long-term studies report that treatment gains are maintained or improved at three years, it is difficult to determine if the effects are actually related to the initial treatment, or, if, in the normal course of migraine, the symptoms have simply remitted.

A third related issue is the possibility of alternative explanations for treatment effects that are not due to the course of migraine itself. It is impossible to control for all historical, maturational, and alternative treatment effects over an extended follow-up period. The possibilities for alternative explanations are nearly endless. A subject may have moved, tried an alternative treatment, married, changed jobs, had children, or had children leave the home. When these issues are not controlled via randomization and controlled comparison groups it is difficult to connect the level of symptomology to a previous treatment.

Finally, selection bias becomes an issue in follow-up studies. Researchers generally only include results of those subjects that can be contacted. It is rare that all subjects who

participated in the initial research study can be contacted. Thus, it is possible that those who are available are significantly different than those who are not. In addition, many researchers complicate the issue of selection bias at follow-up by attempting to contact only those who had a positive outcome to the initial treatment. While intuitively appealing, this procedure selects for not only those who had positive effects, but also for those who can still be reached. This does little to provide information about the true long-term prophylactic effect of a treatment.

These issues suggest that the determination of long-term effects after treatment for migraine prophylaxis is difficult at the best. Fortunately some long-term treatment information is available for up to 5 years on some migraine prophylactic treatments and this information is reviewed below by treatment category.

Propranonol Long-Term Effects

Primary articles included in the present meta-analysis for Propranonol had treatment lengths that ranged from 4-52 weeks. This is the amount of time that the participants were taking Propranonol. Effect size did not significantly correlate with length of treatment (r = -.272), indicating that participants achieved essentially the same reduction in symptoms whether they took Propranonol for 1 month or 12 months. The following authors reported that Propranonol effects were clearly maintained for up to 1 month after discontinuing the medication: Daholf (1987), Kangasniemi et al. (1983), Nadelmann et al. (1986), and Rao et al. (2000). Rao et al. reported that "successful

treatment" effects are maintained for at least 5 months after discontinuation of daily treatment with Propranonol.

Flunarizine Long-Term Effects

Primary articles included in the present meta-analysis for Flunarizine indicate that there is a long-term treatment effect. As previously discussed, Flunarizine shows greater effect sizes the longer it is taken. Studies included ranged in treatment length from 4-24 weeks. Regression analyses indicated that effect sizes were significantly larger when treatment length was longer (R^2 = 5.37, F =15.06). Only one study included in this meta-analysis reported follow-up data on migraine relief after discontinuation of treatment with Flunarizine. Nuti et al. (1996) indicated that follow-up data showed that positive results were maintained for 8.4 months $\pm$ 4.0.

Divalproex Sodium Long-Term Effects

Primary articles included in the present meta-analysis for Divalproex Sodium indicate that effect size does not vary with active treatment length. Treatment length did not significantly correlate with effect size (r = -.223). Treatment lengths varied in the present meta-analysis from 6-24 weeks. None of the articles included in the present meta-analysis provided information on maintenance of effect after discontinuation of active treatment with Divalproex Sodium. Ghose and Niven (1998), in an article that was disqualified from the present meta-analysis, reported that 60% of subjects maintain gains for up 24 months with daily medication. In addition, Rothrock and Mendizable (2000),

also not included in the present meta-analysis, reported that 60% of patients indicated that gains made on Divalproex Sodium were maintained up to 2 months.

Nonpharmacological Long-Term Effects

Thermal biofeedback, relaxation treatments, and combination treatment outcome studies show a different trend than pharmacological treatment studies. The nonpharmacological treatments appear to focus on the effects of making treatments shorter rather than longer. Thermal biofeedback and EMG biofeedback have both experimented with brief therapeutic contacts that utilized at-home practice. The magnitude of the treatment effect sizes *do not* statistically significantly change under these conditions. For example, Blanchard et al. (1985b) reported that outcomes *did not* statistically significantly differ whether they used brief biofeedback/relaxation training (2.6 hours of therapeutic contact) or traditional training (11.6 hours of therapeutic contact).

Articles that were included in the present meta-analysis on thermal biofeedback ranged from 5-32 therapeutic contacts. Relaxation treatments range from 5-12 therapeutic contacts. EMG biofeedback ranged from 5-12 therapeutic contacts. Relaxation treatments and thermal biofeedback plus cognitive therapy ranged from 5-37 weeks and between 4-16 therapeutic contacts. Effect sizes did not vary by treatment length for thermal biofeedback (r = .159) or relaxation therapy (r = .459). Some authors do address how longer term active treatment affects treatment outcomes. Andrasik, Blanchard, Neff, and Rodichok (1984) reported that up to 80% of subjects trained to use progressive muscle relaxation or

progressive muscle relaxation, and thermal biofeedback maintained significant treatment gains at 1 year when given monthly contacts for retraining.

Several reports regarding posttreatment outcomes are available. Many authors report that the effects of thermal biofeedback, relaxation treatments, and EMG treatments are effective for at least 4 weeks after discontinuation of treatment (e.g., Blanchard, Andrisak, Neff, Arena, et al., 1982; Holroyd et al., 1995; Holroyd et al., 1988; Jurish et al., 1983). Some have reported even longer term gains, Blanchard, Appelbaum, Nicholson et al. (1990) reported a 4-month maintenance of treatment gains for thermal biofeedback treatment when the participant achieved an initial 50% reduction in frequency during initial treatment. Daly et al. (1983) reported that participants who were given thermal biofeedback or EMG biofeedback maintained gains for up to three months but that those given relaxation training did not. A six year follow up conducted by Lispers and Ost (1990) indicates that treatment effects are largely maintained for up to 6 years for participants treated with biofeedback (EMG or thermal). Similarly, Sorbi et al. (1989) reported that participants who could be followed for 3 years maintained or increased frequency reductions of migraine when they received relaxation training. Richardson and McGrath (1989) reported positive 6-month outcome effects for participants who received thermal biofeedback and cognitive therapy. They reported that mean treatment frequency scores were statistically significantly lower than baseline at a 6-month follow-up. However, Blanchard et al. (1978) reported that by 4 years those treated with thermal biofeedback or relaxation training were approaching baseline frequencies of migraine headaches.

Summary

The effects of long-term treatment for migraine can be summarized as follows. Once a reduction of symptoms is achieved via any of the above treatments an individual is likely to maintain that reduction of symptoms as long as they are actively being treated. Once the treatment is discontinued the effect may vary by treatment. Propranonol effects may maintain for at least a month, Flunarizine effects may maintain for 8 months, Divalproex Sodium may maintain for up to 2 months, thermal biofeedback and EMG biofeedback may maintain for 6 years, relaxation treatments may maintain for up to 3 years, and thermal biofeedback plus cognitive treatments may maintain for 6 months. However, all of these results should be interpreted with caution due to methodological problems inherent in the studies.

Practical Implications

The findings of the present meta-analysis do little to clarify which treatment should be recommended. The results indicate that most of the treatments reviewed produce nearly the same effect size. All effects (excluding those from EMG biofeedback treatments) are considered moderate to large. Thus, individuals who choose from among Propranonol, Flunarizine, Divalproex Sodium, thermal biofeedback, relaxation therapy, and thermal biofeedback plus cognitive therapy for migraine prophylaxis are likely to see some reduction of symptoms.

As a result determining the treatment of choice must take into account factors other than effectiveness rates. Ward (2000) suggested that the selection of treatment for migraine prophylaxis should be made on the basis of comorbid disorders. If another condition exists that can simultaneously be treated with one of the agents, then this treatment should be selected first. Ward reported that some commonly found conditions (and treating agents) that should be considered are hypertension (Propranonol), mitral valve prolapse (Propranonol), anxiety (Propranonol), bipolar disorder (Divalpoex Sodium), epilepsy (Divalproex Sodium), Raynauds disease (Flunarizine). An excellent decision tree that considers comorbid disorders and contraindications for migraine prophylaxis pharmacological treatment appears in Adelman and Von Seggern (1995).

Ward (2000) did not mention the use of nonpharmacological treatments but the same logic applies. Previous research shows that these nonpharmacological treatments work well for conditions other than migraine prophylaxis. Researchers have documented that thermal biofeedback and relaxation therapy are effective in treating anxiety (Culpepper, 2002), depression (Setter & Kupper, 2002), and high blood pressure (Setter & Kupper). Thermal biofeedback in particular has been noted to treat circulation issues in Raynauds disease (Sappington & Fiorito, 1985). As can be seen, Ward's suggestion to select via comorbid disorders narrowed the fields of treatments, but did not identify a single best choice treatment.

Another consideration in choosing a treatment may be compliance. As previously mentioned compliance with treatment regimes is a documented problem in migraine prophylaxis (Mullerners et al., 1998). The above-researched techniques will not have the

complete desired effect if the individual is unable to comply with treatment recommendations. If compliance issues are of concern for an individual, then issues that impact compliance should be considered when selecting the treatment of choice. One way of selecting the best treatment may be choosing the one that requires the least time and energy expenditure. Of the above treatments, Propranonol and Flunarizine are likely to be adhered to with the least amount of problems (each treatment can be administered on a once-a-day schedule). Given that Flunarizine is not available in the United States, Propranonol would be the treatment of choice. Divalproex Sodium requires greater compliance because it may be administered up to three times a day and requires regular laboratory exams for safety reasons. All nonpharmacological treatments require more time investment from the treated individual. However, an advantage to using nonpharmacological treatment is the ability of the clinician to monitor compliance. That is, many treatments involve components with the individual participating in their treatment in the presence of the clinician.

When selecting the treatment of choice patient preference cannot be ignored. The side effects profiles, time commitments, and long-term effects vary widely between these treatments. Individuals often have specific belief systems that influence their preferences for treatment. Individuals needing migraine prophylactic treatment should be informed of the options and included in the decision making about treatment of choice so they can be active participants in their own health care (Capobianco et al., 1996). Individuals who are involved in making their own treatment decisions are more likely to value the treatment they have chosen and, thus, adhere better to treatment guidelines.

Finally, cost considerations need to be addressed. Adelman and Von Seggern (1995) identified that Propranonol (the twice-a-day preparation) as the least expensive treatment. However, the one-time-a-day preparation was almost four times as expensive. Divalpoex Sodium weighed in at almost 12 times the cost of the twice-a-day Propranonol. Estimates for Flunarizine are not available given that it is not legal to sell within the United States. The above estimates include only medication costs. The cost of doctor visits, laboratory work, and third-party fees are not estimated. So this is an estimate of only a part of the costs associated with the pharmacological treatments. Osterhaus and Townsend (1991) estimated at that time that other medical costs per year per patient for migraine management included $281 for emergency room visits, $148 for clinic visits, and $387 for hospitalizations, averaging an additional $68 a month for pharmacological treatments.

Nonpharmacological treatments are initially expensive. Treatment per session prices can average $80. If seen weekly, the cost per month could be $240, where Divalproex Sodium plus monthly medical expenses are estimated at $158 per month. Attanasio, Andrasik, and Blanchard (1987) reported that there was little difference in the cost effectiveness between nonpharmacological treatments. All nonpharmacological treatments have similar costs associated with the initial training. Importantly, Blanchard, Jaccard, Andrasik, Guarnieri, & Jurish (1985) report that nonpharmacological treatments significantly reduced costs after termination of treatment due to long-term effects and patient self-management. As previously mentioned, long-term studies evaluating the effects of nonpharmacological treatment effects indicated that reduced frequency of

migraine is maintained up to 6 years (Lisspers & Ost, 1990). Pharmacological treatments have very little support as to having a lasting treatment effect beyond a few months after the active treatment phase. This suggests the costs associated with nonpharmacological treatments are one-time costs where pharmacological costs are ongoing. Blanchard, Jaccard, et al. (1985) reported that medical costs associated with treating headaches were approximately $1,000 per year prior to nonpharmacological treatment and approximately $50 per year following treatment (costs include medical costs and lost productivity). The long term effects after treatment is discontinued indicate that nonpharmacological treatments may be more cost effective than pharmacological treatments in the long run.

In sum, a variety of factors influence treatment choice for migraine prophylaxis. There is no single best choice for all individuals. Comorbid disorders, cost, and compliance are just a few of the considerations in making the choice. However, the current meta-analysis suggests that once the choice is made, if compliance is high, the treatment is more likely than placebo to reduce frequency of headaches.

Limitations

The current findings are limited in several ways. First, these findings are directly reliant on the quality and accuracy of the findings in the primary research. Second, the combination of these two major bodies of research has resulted in the mixing of some groups that may not be compatible. Finally, significant problems exist in the primary literature for nonpharmacological treatments that limit their findings. Each of these limitations are discussed below.

Quality of Primary Research

These findings cannot outstretch the initial accuracy of the articles on which they are based. The current meta-analysis attempted to address this issue through the use of quality ratings. However, it should be noted that the quality scores used in this meta-analysis only address a few factors that could compromise outcome validity and reliability. Thus, other serious threats or flaws may exist in the primary research that would challenge the findings of this meta-analysis. The present meta-analysis supports the findings reported by other meta-analytic studies (Duke University and The Center for Clinical Health and Policy, 1999a, 1999b; Holroyd & Penzien; 1990). Nevertheless, all of these meta-analyses have been based on the same primary research literature and may be subject to the same primary flaws.

Summing Data

Subjecting the two bodies of research (pharmacological and nonpharmacological) to the same standards and coding methods revealed several differences between the two literature bases that may limit these findings. There are several procedural and nomenclature differences between the two groups that may cause the outcome findings to be limited. For example, the term "vascular headache" is one commonly used in the behavioral research. This term is defined as some form of migraine (mixed, classic, or common). Researchers in the behavioral area often report results for the vascular headache subjects as a group. The pharmacological research does not typically use mixed groups. They tend to select subjects who have very specific migraine symptoms and then report

results in two migraine categories (common or classic). The research included in this meta-analysis for pharmacological treatments rarely included mixed headache patients. However, subjects with transformed headache (a condition where individuals with migraine headaches begin to have daily tension headaches intermixed with migraines due to medication/reinforcement issues) are included in some pharmacological studies. Subjects with transformed headache are not addressed in the nonpharmacological research. This meta-analysis combines this diverse information into single summary effect sizes for migraine headache treatments. In the process of summing these two diverse bodies of literature some specific information may have been lost. Thus, the comparison of the outcome of the treatments is limited to broad generalizations.

Problems in the Nonpharmacological Literature

The findings for nonpharmacological treatments are significantly limited by (a) the lack of studies that use an appropriate control or attention placebo group, and (b) an overreliance on the percent improved score. The paucity of published controlled studies on nonpharmacological treatments is not from the lack of trying. Rather, studies designed as controlled trials have often selected attention placebo conditions that later demonstrate an active effect. Researchers have attempted to use relaxation (Lacroix et al., 1983), pseudo-meditation (Blanchard, Appelbaum, Radnitz, et al.. 1990), hypnosis (Reich, 1989), false biofeedback (Reading, 1984), home biofeedback (Blanchard et al., 1985a), and contact with a therapist by phone (Richardson & McGraph, 1989) as attention placebos. However, each of these groups has shown similar effects to the active treatment group.

The only control groups that appeared to consistently not have big treatment gains were wait list groups. This phenomenon resulted in markedly fewer controlled studies that were available for systematic review for nonpharmacological treatment than for pharmacological treatments.

The second limiting factor found in the existing literature on nonpharmacological treatments was the overreliance on the percent improved score. The percent improved score was the individual improvement scores average for the *group* of treated individuals. For example, authors may report that those in the thermal biofeedback group had a percent improved score of 48%. Many authors in the nonpharmacological migraine treatment area report this score as their primary outcome index. When this is the case, individual data are lost because it is impossible to tell if the individuals in the treatment group went from 4 headaches a month to 2 headaches a month, or from 10 headaches a month to 5 headaches a month. Both are significant changes but have different meanings. In addition, it is not standard to report the percent improved score with variance scores. Therefore, it is impossible to tell if most subjects in the study were near the reported percent improved score or if there was a great deal of deviation. Without standard reporting of variance these scores become less meaningful.

The combination of the lack of a true control condition and inadequate statistical reporting (primarily due to the overuse of the percent improved score) resulted in a large number of primary research articles that could not be quantitatively compared. These studies did not lend themselves to standardization via quantitative methods and thus had to be disqualified from the current meta-analysis. As a result, this meta-analysis is based on

only a fraction of primary articles available on nonpharmacological treatments. Therefore, the conclusions drawn from these findings may be limited by being based on only a small portion of the available research findings.

Suggestions for Future Research

Continued research in the area of migraine prophylaxis is recommended, however, prior to beginning new trials researchers need to recognize what is *not* needed in this area. Research on why biofeedback treatments are superior to relaxation treatments for migraine prophylaxis is no longer needed because the current body of literature shows that biofeedback is not superior to relaxation treatments. Though not directly addressed in this meta-analysis there is also no current need to establish the ideal body temperature change that a person needs to obtain for biofeedback to be effective. Results clearly indicate that classic and common migraine patients cannot change their body temperature as readily as other type of subjects; however, ability to reach certain temperature changes does not correlate with the decrease of headache frequency (Werbach & Sandweiss, 1978). Another course of action that no longer seems warranted is research on combination therapies that do not use two empirically validated treatments. There is not a valid rationale for throwing treatments together that have not been established as effective by themselves (e.g., EMG biofeedback plus hypnosis with autogenic training). Finally, the role of Propranolol is clearly established in migraine prophylaxis. There is little need for continued controlled trials to replicate this finding.

These areas aside, the findings of present meta-analysis indicate future research is needed in the area of migraine prophylaxis. More research is needed on (a) controlled trials on nonpharmacological treatments, (b) the effect of combining empirically supported treatments from the two categories (nonpharmacological and pharmacological), and (c) the development of a comprehensive theory on migraine that combines findings from different fields. Each of these suggestions are discussed below.

First, the value of nonpharmacological prophylactic treatments (other than thermal biofeedback) needs to be clearly established. This can only be done through clearly controlled basic research studies that replicate outcome findings. The lack of basic research on commonly accepted techniques for nonpharmacological treatments is alarming. Many authors have shifted on to investigating correlates of outcomes for relaxation and cognitive techniques, citing as part of their study rational that the benefits of these treatments are clearly documented by other authors (e.g., Hart, 1984). This is not a valid rationale for all nonpharmacological treatments. As shown in the current meta-analysis, there is a large body of research regarding the use of these techniques for migraine prophylaxis; however, very little of the research works to establish that the basic effects of these treatments are greater than placebo/control group gains.

A related issue is the need for the identification of an appropriate attention placebo condition for migraine prophylaxis. The definition of attention placebo as it relates to behavioral interventions has often been debated. Definitions for attention placebo as it applies to nonpharmacological treatments include: (a) a treatment or procedure that has no specific effect, but that can be presumed to have an effect (Shapiro & Morris, 1978), (b) a

substitute for a genuine treatment (Senger, 1987), and (c) nonspecific treatment facters that encourage hope, learning, sharing, emotional arousal, and mastery, but do not provide any specific effect (Fish, 1973). The variety of definitions available for attention placebo stems largely from the attempt to apply a medical term to psychotherapy. The term placebo in medical arena's refers to the procedure of giving an inactive substance in the place of a chemically active substance to rule out interpersonal and expectation factors as a cause of outcome (Senger). This, obviously, is a difficult concept to apply to nonpharmacological interventions that rely on interpersonal factors for treatment. This has led some authors to suggest that effects of psychological and behavioral treatments are nothing more than extensive placebo effects (e.g., Patterson, 1967). Thus, it may be argued that the observed effects for nonpharmacological prophylactic treatments for migraine are merely attention placebo effects. A strong argument against this comes from Mathew (1981), who reported multi-arm treatment findings. He suggested that thermal biofeedback and Propranolol have similar effectiveness rates, but that both are better than placebo. If thermal biofeedback were just a strong placebo effect, one would expect the efficacy rates to be lower than Propranolol and more equivalent to placebo results. However, this is not the case.

Thus, it is likely that thermal biofeedback, relaxation therapy, and combined therapies have a specific effect that causes the reduction of migraine frequency. If this is the case, this effect needs to be demonstrated above and beyond hope, expectation, and interpersonal factors through controlled studies that employ attention placebo groups. In the pharmacological migraine prophylaxis literature, the placebo effect has been

demonstrated to account for at least 35% improvement in migraine frequency (Couch, 1987). If this same approximate rate can be applied to nonpharmacological treatments, then an attention placebo could be defined as a treatment that does not include specific treatment factors that causes approximately 35% improvement but does not equal total treatment effect of the active treatment. The present study will use this definition of attention placebo.

The difference between attention placebo groups and active treatment groups needs to be established in the nonpharmacological literature. It is apparent in the existing body of literature on nonpharmacological migraine prophylaxis that any attention placebo condition that involves regular visits, sitting alone in a treatment room, machines that have face validity, or guided relaxation treatment at home, will show a treatment effect and, therefore, not be an attention placebo. This suggests that the nonpharmacological treatments have a common active component that has yet to be explained. It is possible that the act of routinely engaging in a quieting or calming activity or that the education provided when the treatments are undertaken, or that providing a rationale to participants may be the common active feature. Many authors have documented the fact that the nonpharmacological treatments are better than wait list controls (e.g., Blanchard, Appelbaum, Nicholson, et al., 1990; Blanchard, Nicholson, Radnitz, et al., 1991; Gauthier, et al., 1985). It can then be assumed that the nonpharmacological treatments are providing treatment beyond time and daily self- monitoring, but how they are providing this treatment is still uncertain. An important part of discovering the active components in nonpharmacological treatment will be identifying an attention placebo that does not elicit

the same treatment effect. The studies included in the current meta-analysis in the EMG biofeedback plus thermal biofeedback plus relaxation treatment condition may offer some guidance in this area. This treatment category included three primary articles that used control groups that document approximately 20-30% gains. These gains are similar to those often reported for placebo (i.e., 35%). Two of the studies included control groups who were told to go home and relax without specific instruction (McGrady et al., 1994; Waquier et al., 1995). The fact, that the authors were able to document treatment effects for EMG biofeedback plus thermal biofeedback plus relaxation above and beyond their control group is promising. It should be noted that the net result in the current meta-analysis of using control conditions rather than a wait list group was a lower summary effect size for this treatment (ln (OR) = 1.11). This may suggest that when a true attention placebo for nonpharmacological migraine prophylactic treatment outcome studies is found, the noted treatment effect sizes for the nonpharmacological treatments may be reduced (due to the methods of calculating effect sizes)

It appears that adding treatments within a category (nonpharmacological or pharmacological) does little to improve the efficacy of the treatment. This is most clearly see in the nonpharmacological treatments. Results show that most treatments in the nonpharmacological area have the same effect even if they are combined (e.g., thermal biofeedback and thermal biofeedback plus relaxation plus cognitive therapy have nearly the same effect sizes). However, future research in the area of cross-category combination treatment is needed. Taking a treatment from the nonpharmacological category and one from the pharmacological category that are both empirically supported and combining

them is a relatively unexplored area. There are currently only two articles that examine the effects of one of these combinations; thermal biofeedback plus Propranonol (Holroyd et al., 1995; Mathew, 1981) Both studies reported promising results. Holroyd, France, et al. reported that the use of a combined thermal biofeedback and Propranonol approach was significantly better than using biofeedback alone. Mathew (1981) reported that thermal biofeedback plus Propranonol was superior to Propranonol alone, or thermal biofeedback alone. In addition, Mathew reported that biofeedback plus Propranonol was superior for migraine prophylaxis to all seven treatments examined in the same article (control, Propranonol, Amitriptyline, thermal biofeedback, Propranonol plus Amitriptyline, Amitriptyline plus thermal biofeedback, and Propranonol plus Amitriptyline plus thermal biofeedback). However, the author notes that these results are for migraine headache only and that results for mixed headache differ. It is interesting to note that both articles are relatively old by research standards and that little has been done to follow-up on their findings.

A primary need in the area of migraine research is a new guiding theory that takes into account the available research findings. Much of the research on migraine prophylaxis appears haphazard and undirected. This may be due to the lack of a unifying theory about the course of migraines and how variables interact with one another. There seems to be very little work being published that attempts to integrate findings on migraine from various research fields. Pharmacological, nonpharmacological, genetic, and biological researchers are reporting some perplexing findings that need to be addressed. For example, why do false biofeedback and true biofeedback not statistically significantly differ

in outcome for migraine prophylaxis when they both are statistically better than wait list groups? Or why does Propranonol seem to be one of the few beta-blockers that actually works as a migraine prophylactic? And why do Propranonol, thermal biofeedback, or practicing home relaxation seem to all have similar effect?

The development of a theoretical hypothesis that synthesizes these and many other results could direct researchers toward potentially fruitful areas of research and away from repeated problems of the past. The unification of many fields through theory would allow researchers to work with each other. Potentially, a concrete theory could help eliminate the irrelevant and redundant research that is present in this body of literature (e.g., 30 years of research that uses relaxation as a control group or repeated studies on the brand of machine needed to give biofeedback when a basic thermometer seems to work just as well).

In sum, research on migraine and migraine prophylaxis needs to be theory driven. Ideally, the theory would address the pathogenesis, genetic research, psychological contributions, common pathways, and treatment of migraine. The theory would address the results of the present meta-analysis and the results of other meta-analytic studies that indicate treatments have similar effect sizes and none of these effect sizes are as high as one would hope after 30 years of research. In addition, the theory would suggest specific hypothesis that could be tested in various fields. Such a unifying theory seems crucial to reaching a better understanding of the complexity of migraine headaches.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

Abrami, P.C., Cohen, P.A., & d'Apollonia, S. (1988). Implementation problems in meta-analysis. *Review of Educational Research, 58*, 155-179.

Adelman, J.U., & Von Seggern, R. (1995). Cost considerations in headache treatment. Part 1: Prophylactic migraine treatment. *Headache, 35,* 479-487.

*Ahuja, G.K., & Verma, A.K. (1985). Propranolol in prophylaxis of migraine. *Indian Journal of Medical Research, 82*, 263-265.

*Albers, G.W., Simon, L.T., Hamik, A., & Peroutka, S.J. (1989). Nifedipine versus propranolol for the intial prophylaxis of migraine. *Headache, 29,* 214-217.

*Al Deeb, S.M., Biary, N., Bahou, Y., Al Jaberi, M., & Khoja, W. (1998). Flunarizine in migraine: A double-blind placebo-controlled study (in a Saudi population). *Headache, 32,* 461-462.

Al-Qassab, H.K., & Findley, L.J. (1993). Comparison of Propranolol LA 80 mg and Propranolol 160 mg in migraine prophylaxis: A placebo controlled study. *Cephalagia, 13,* 128-131.

Alder, C.S., & Alder, S.M. (1976). Biofeedback-psychotherapy for the treatment of headaches: A 5-year follow-up. *Headache, 16,* 189-191.

*Allais, G., De Lorenzo, C., Quirico, P.E., Airola, G., Tolardo, G., Mana, O., & Benedetto, B. (2002). Acupuncture in the prophylactic treatment of migraine without aura: acomparison with flunarizine. *Headache, 42,* 855-861.

Amery, W.K. (1983). Flunarizine, a calcium channel blocker: A new prophylactic drug in migraine headache. *Headache, 23,* 70-74.

Amery, W.K., Wauquier, A., Van Nueten, J.M., De Clerck, F., Van Reempts, J.V., & Janssen, P.A.J. (1981). The anti-migrainous pharmacology of flunarizine (R 14 950), a calcium antagonist. *Drugs and Experimental Clinical Research, VII,* 1-10.

Anderson, J.A., Basker, M.A., & Dalton, R. (1975). Migraine and hypnotherapy. *International Journal of Clinical and Experimental Hypnosis, 23,* 48-58.

Andersson, P.G. (1985). Migraine in patients with cluster headache. *Cephalagia, 5,* 11-16.

*Andersson, P.G. & Petersen, E.N. (1981). Propranolol and femoxetine, a 5HT-uptake inhibitor, in migraine prophylaxis. *Acta Neurologica Scandinavia, 64,* 280-288.

Andrasik, F., Blanchard, E.B., Neff, D.F., & Rodichok, L.D. (1984). Biofeedback and relaxation training for chronic headache: A controlled comparison of booster treatments and regular contacts for long-term maintenance. *Journal of Consult Clinical Psychology, 52,* 609-615.

Andrasik, F., Pallmeyer, T.P., Blanchard, E.B., & Attanasio, V. (1984). Continuous versus interrupted schedules of thermal biofeedback: An exploratory analysis with clinical subjects. *Biofeedback and Self Regulation, 9,* 291-298.

Andreychuk, T., & Skriver, C. (1975). Hypnosis and biofeedback in the treatment of migraine headache. *International Journal Clinical and Experimental Hypnosis, 23,* 72-83.

Angoli, A., Bussone, G., Mailland, F., Manzoni, G.C., Martucci, N., & Nappi, G. (1991). Dihydroergokryptine vs. Flunarizine in the basic treatment of migraine without aura. *Cephalagia, 11,* 216-217.

Attanasio, V., Andrasik, F. & Blanchard, E.B. (1987). Cognitive therapy and relaxation training in muscle contraction headache: Efficacy and cost-effectiveness. *Headache, 27,* 254-260.

Baker, C. (1987). Double-blind evaluation of flunarizine and placebo in the prophylactic treatment of migraine. *Headache, 27,* 288.

Baldrati, A., Cortelli, P., Procaccianti, G., Gamberini, G., D'Alessandro, R., Baruzzi,A., & Sacquegna, T. (1983). Propranolol and acetylsalicylic acid in migraine prophylaxis. Double-blind-crossover study. *Acta Neurologica Scandinavia, 67,* 181-186.

Bassi, P., Brunati, L., Rapuzzi, B., Alberti, E., & Mangoni, A. (1992). Low dose flunarizine in the prophylaxis of migraine. *Headache, 32,* 390-392.

Behan, P.O., & Reid, M. (1980). Propranolol in the treatment of migraine. *Practitioner, 24,* 1-3.

Billings, R.F., Thomas, M.R., Rapp, M.S., Reyes, E., & Leith, M. (1984). Differential efficacy of biofeedback in headache. *Headache, 24,* 211-215.

Blanchard, E.B. (1987). Long-term effects of behavioral treatment of chronic headache. *Behavior Therapy, 18,* 375-385.

Blanchard, E.B. (1994). Behavioral Medicine and Health Psychology. In A.E. Bergin & S.L. Garfield (Eds.), *Handbook of Psychcotherapy and Behavioral Change-Fourth Edition.* New York: Wiley.

Blanchard, E.B., Andrasik, F., Ahles, T.A., Teders, S.J., & O'Keefe, D. (1980). Migraine and tension headache: A meta-analytic review. *Behavior Therapy, 11,* 613-3-631.

*Blanchard, E.B., Andrasik, F., Appelbaum, K.A., Evans, D.D., Jurish, S.E., Teders, S.J., Rodichok, L.D., & Barron, K.D. (1985a). The efficacy and cost-effectiveness of minimal-therapist-contact, non-drug treatments of chronic migraine and tension headache. *Headache, 25,* 214-220.

*Blanchard, E.B., Andrasik, F., Appelbaum, K.A., Evans, D.D., Jurish, S.E., Teders, S.J., Rodichok, L.D., & Barron, K.D. (1985b). The efficacy and cost-effectiveness of minimal -therapist-contact non-drug treatment of chronic migraine and tension headache. *Headache, 25,* 214-220.

*Blanchard, E.B., Andrasik, F., Evans, D.D., Neff, D. F., Appelbaum, K.A., & Rodichok, L.D. (1985c). Behavioral treatment of 250 chronic cheadache patients: A clinical replication series. *Behavioral Therapy, 23,* 308-327.

*Blanchard, E.B., Andrasik, F., Neff, D.F., Arena, J.G., Ahles, T.A., Jurish, S.E., Pallmeyer, T.P., Saunders, N.L., Teders, S.J., Barron, K.D., & Rodichok, L.D. (1982). Biofeedback and relaxation training with three kinds of headache: Treatment effects and their prediction. *Journal of Consulting and Clinical Psychology, 50,* 562-575.

Blanchard, E.B., Andrasik, F., Neff, D.F., Teders, S.J., Pallmeyer, T.P., Arena, J.G., Jurish, S.E., Saunders, N.L., Ahles, T.A., & Rodichok, L.D. (1982). Sequential comparisons of relaxation training and biofeedback in the treatment of three kinds of chronic headache or, the machines may be necessary some of the time. *Behavior Research Therapy, 20,* 469-481.

Blanchard, E.B., Appelbaum, K.A., Guarnieri, P., Morrill, B., & Dentinger, M.P. (1987). Five year prospective follow-up on the treatment of chronic headache with biofeedback and/or relaxation. *Headache, 27,* 580-583.

*Blanchard, E.B., Appelbaum, K.A., Nicholson, N.L., Radnitz, C.L., Morrill, B., Michultka, D., Kirsch, C., Hillhouse, J., & Dentinger, M.P. (1990). A controlled evaluation of the addition of cognitive therapy to a home-based biofeedback and relaxation treatment of vascular headache. *Headache, 30,* 371-376.

*Blanchard, E.B., Appelbaum, K.A., Radnitz, C.L., Morrill, B., Michultka, D., Kirsch, C., Guarnieri, P., Hillhouse, J., Jaccard, J., & Barron, K.D. (1990). A controlled evaluation of thermal biofeedback and thermal biofeedback combined with cognitive therapy in the treatment of vascular headache. *Journal of Consulting and Clinical Psychology, 58,* 216-224.

Blanchard, E.B., Jaccard, J., Andrasik, F., Guarnieri, P., & Jurish, S.E. (1985). Reduction in headache patients' medical expenses associated with biofeedback and relaxation treatments. *Biofeedback and Self-Regulation, 10,* 63-68.

*Blanchard, E.B., Nicholson, N.L., Radnitz, C.L., Steffek, B.D., Appelbaum, K.A., & Dentinger, M.P. (1991). The role of home practice in thermal biofeedback. *Journal of Consulting and Clinical Psychology, 59,* 507-512.

*Blanchard, E.B., Peters, M.L., Hermann, C., Turner, S.M., Buckley, T.C., Barton, K., & Dentinger, M.P. (1997). Direction of temperature control in the thermal biofeedback treatment of vascular headache. *Applied Psychophysiology Biofeedback, 22,* 227-245.

*Blanchard, E.B., Theobald, D.E., Williamson, D.A., Silver, B.V., & Brown, D.A. (1978). Temperature bofeedback in the treatment of migraine headaches. *Archives of General Psychiatry, 35,* 581-588.

Blau, J.N. (1980). Migraine prodromes separated from the aura: Complete migraine. *British Journal of Medicine, 281,* 658-660.

Bogaards, M.C. & ter Kuile, M.M. (1994). Treatment of recurrent tension headache: A meta-analytic review. *Clinical Journal of Pain, 10,* 174-190.

Bonuso, S., Di Stasio, E., Marano, E., Sorge, F., & Leo, A. (1986). Sublingual flunarizine: a new effective management of the migraine attack. A comparison versus ergotamine. *Headache, 26,* 227-230.

Bordini, C.A., Arruda, M.A., Ciciarelli, M.C., & Speciali, J.G. (1997). Propranolol and Flunarizine vs. Flunarizine plus Propranolol in migraine without aura prophylaxis. *Arquivos de neuro-psiquiatria, 55,* 536-541.

*Borgesen, S.E., Neilsen, J.L., & Moller, C.E. (1974). Prophylactic treatment of migraine with propranolol. *Acta Neurologica Scandinavia, 50,* 651-656.

Brockwell, S.E., & Gordon, I.R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine, 20,* 825-840.

Brown, J.M. (1984). Imagery coping strategies in the treatment of migraine. *Pain, 18,* 157-167.

*Bussone, G., Baldini, S., D'Andrea, G., Cananzi, A., Frediani, F., Caresia, L., Ferro Milone, F., & Boiardi, A. (1987) Nimodipine versus flunarizine in common migraine: A controlled pilot trial. *Headache, 27,* 76-79.

Caroll, D. (1996). Evidence-based pain management: The need for change. *Canadian Journal of Nursing Research, 28,* 59-66.

Carroll, J.D., Reidy, M., Savundra, P.A., Cleave, N., & McAish, J. (1990). Long ancting propranolol in the prophylaxis of migraine: A comparative study of two doses. *Cephalagia, 10,* 101-105.

Capobianco, D.J., Chesire, W.P., & Campbell, J.K. (1996). An overview of the diagnosis and pharmacologic treatment of migraine. *Mayo Clinic Procedures, 71,* 1055-1066.

Centonze, V., Attolini, E., Macinagrossa, G., Campanozzi, F., Magrone, D., Tesauro, P., Loisy, C., & Albano, O. (1985). 5 mg. flunarizine in preventive therapy of migraine: Efficacy and tolerance. *Cephalagia, 5*(Suppl. 3), 536-537.

Centonze, V., Magrone, D., & Vino, M. (1990). Flunarizine in migraine prophylaxis: Efficacy and tolerability of 5 mg. and 10 mg. dose levels. *Cephalagia, 10,* 17-24.

Cerbo, R., Casacchia, M., Formisano, R., Feliciani, M., Cusimano, G., Buzzi, M.G., & Agnoli, A. (1986). Flunarizine-pizotifen single-dose double-blind cross-over trial in migraine prophylaxis. *Cephalagia, 6,* 15-18.

Chambless, D. L., & Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7-18.

Chang, B., Waternaux, C., & Lipsitz, S. (2001). Meta-analysis of binary data: Which within-study variance estimate to use? *Statistics in Medicine, 20,* 1947-1956.

Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society, 4,* 102-118.

Cohen, M.J., McArthur, D.L., & Rickles, W.H. (1980). Comparison of four biofeedback treatments for migraine headache: physiological and headache variables. *Psychosomatic Medicine, 42,* 463-480.

Cortelli, P., Sacquegma. T., Albani, F., Baldrati, A., D'Alessandro, R., Baruzzi, A., & Lugaresi, E. (1985). Propranolol plasma levels and the relief of migraine: Relationship between plasma propranolol and 4-Hydorxypropranolol concentrations and clinical effectiveness. *Archives of Neurology, 42*, 46-48.

Compas, B.E., Haaga, D.A.F., Keefe, F.J., Leintenberg, H., & Williams, D.A. (1998). Sampling of empirically supported psychological treatment for health psychology: Smoking, chronic pain, cancer, and bulimia nervosa. *Journal of Consulting and Clinical Psychology, 66,* 89-112.

Couch, J.R., Jr.(1987) Placebo effect and clinical trials in migraine therapy. *Neuroepidemiology, 6,* 178-185.

Culpepper, L. (2002). Generalized anxiety disorder in primary care: Emerging issues in management and treatment. *Journal of Clinical Psychiatry, 63*(Suppl. 8), 35-42.

Curtin, F., Altman, D.G., & Elbourne, D. (2002). Meta-analysis combining parallel and cross-over trials. I: Continuous outcomes. *Statistics in Medicine, 21,* 2131-2114.

*Daholf, C. (1987). No clearcut long-term prophylactic effect of one month treatment of propranolol with migraineurs. *Cephalagia, 7*(Suppl. 6), 459-460.

Dalessio, D.J., (1994). Diagnosing the severe headache. *Neurology, 44,* S6-S12.

*Daly, E.J., Donn, P.A., Galliher, M.J., & Zimmerman, J.S. (1983). Biofeedback applications to migraine and tension headaches: A double-blinded outcome study. *Biofeedback and Self Regulation, 8,* 135-152.

D'Amato, C.C., D'Amato, A.C., Alfano, V., Giodano, E., & Marmo, E. (1990). Flunarizine in long-term migraine prophylaxis: Clinical evidence. *Journal of Medicine, 21,* 201-207.

de Lissovoy, G., & Lazarus, S.S. (1994). The economic cost of migraine. Present state of knowledge. *Neurology, 44*(6)(Suppl. 4), S56-62.

Deeks, J.J. (2002). Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine, 21,* 1575-1600.

Diamond, S., Kudrow, L., Stevens, J., & Shapiro, D.B. (1982). Long-term study of propranolol in the treatment of migraine. *Headache, 22,* 268-271.

*Diamond, S. & Medina, J.L. (1976). Double blind study of propranolol for migraine prophylaxis. *Headache, 16,* 24-26.

Diamond, S., & Montrose, D. (1984). The value of biofeedback in the treatment of chronic headache: A four-year retrospective study. *Headache, 24,* 5-18.

Diamond, S., & Schenbaum, H. (1983). Flunarizine, a calcium channel blocker, in the prophylactic treatment of migraine. *Headache, 23,* 39-42.

*Diener, H.C., Foh, M., Iaccarino, C., Wessely, P., Isler, H., Strenge, H., Fishcer, M., Wedekind, W., & Taneri, Z.- on behalf of the study group. (1996). Cyclandelate in the propylaxis of migraine: A randomized, parallel, double-blind study in comparison with placebo and propranolol. *Cephalagia, 16*, 441-447.

Diener, H.C., Matias-Guiu, J., Hartung, E., Pfaffenrath, V., Ludin, H.P., Nappi, G., & DeBeukelaar, F. (2002). Efficacy and tolerability in migraine prophylaxis of flunarizine in reduceddoses: A comparison with propranolol 160 mg daily. *Cephalagia, 22,* 209-21.

Duke University and The Center for Clinical Health Policy Research. (1999a). *Behavioral and physical treatments for migraine.* Rockville, MD: Agency for Health Care Policy and Research. (NTIS No. PB99-127953-2.2)

Duke University and The Center for Clinical Health Policy Research. (1999b). *Drug treatments for the prevention of migraine headache*. Rockville, MD: Agency for Health Care Policy and Research. (NTIS No. PB99-127953-2.3)

Dunlop, W.P., Cortina, J.M., Vaslow, J.B., & Burke, M.J. (1996). Meta-analysis of experiments with matched groups or repeated measure designs. *Psychological Methods, 1,* 170-177.

Elbourne, D.R., Altman, D.G., Higgins, J.P.T., Curtin, F., Worthington, H.V., & Vail, A. (2002). Meta-analyses involving cross-over trails: Methodolgical issues. *International Journal of Epidemiolology, 31,* 140-149.

Fahrion, S.L. (1977). Autogenic biofeedback treatment for migraine. *Mayo Clinic Proceedings, 52,* 76-84.

Featherstone, H. J., (1985). Migraine and muscle contraction headaches: A continuum. *Headache, 25*, 194-198.

Feuerstein, M., & Adams, H.E. (1977). Cephalic vasomotor feedback in the modification of migraine headache. *Biofeedback and Self Regulation, 2,* 241-254.

Feuerstein, M., Bortolussi, L., Houle, M., & Labbe, E. (1983). Stress temporal artery activity, and pain in migraine headache: A prospective analysis. *Headache, 23,* 296-304.

Fish, J.M. (1973). *Placebo therapy.* San Francisco: Jossey Bass.

Flor, H., Fydrich, T., & Turk, D.C. (1992). Efficacy of multidisciplinary pain treatment centers: A meta-analytic review. *Pain, 49,* 221-230.

Ford, M.R., Stroebel, C.F., Strong, P., & Szarek, B.L. (1983). Quieting response training: Predictors of long-term outcome. *Biofeedback and Self Regulation, 8,* 393-408.

*Forssman, B., Henriksson, K., Johannsson, V., Lindvall, L., & Lundin, H. (1976). Propranolol for migraine prophylaxis. *Headache, 16,* 238-245.

*Freitag, F. G., Diamond, S., & Diamond, M. (1991). A placebo controlled trial of flunarizine in migraine prophylaxis. *Cephalagia, 11*, 157-158.

*French, D.J., Gauthier, J., Roberge, C., Bouchard, S., & Nowen, A. (1997). Self-efficacy in the thermal biofeedback treatment of migraine headache. *Behavior Therapy, 28*(1), 109-125.

*Frenken, C.W., & Nuijten, S.T.M. (1984). Flunarizine, a new preventive approach to migraine. *Neurological Neurosurgery, 86*, 17-20.

Friedman, H., & Taub, H.A. (1984). Brief psychological training procedures in migraine treatment. *American Journal of Clinical Hypnosis, 26,* 187-200.

Gainer, J.C. (1978). Temperature discrimination training in biofeedback treatment of migraine. *Journal of Behavior Therapy and Experimental Psychiatry, 9,* 185-187.

Gall, M.D., Borg, W.R., & Gall, J.P. (1996). *Educational research. An introduction* (6th ed.). White Plains, NY: Longman.

Gallagher, R.M., & Warner, J.B. (1984). Patient motivation in the treatment of migraine. A nonmedicinal study. *Headache, 24,* 269-271.

Gamble, E.H., & Elder, S.T. (1983). Multimodal biofeedback in the treatment of migraine. *Biofeedback and Self Regulation, 8,* 383-392.

Gauthier, J., Bois, R., Allaire, D., & Drolet, M. (1981). Evaluation of skin temperature biofeedback and training at two different sites for migraine. *Journal of Behavioral Medicine, 4,* 407-419.

Gauthier, J., & Carrier, S. (1991). Long-term effects of biofeedback on migraine headache: A prospective follow-up study. *Headache, 31,* 605-612.

*Gauthier, J., Cote, G., & French, D. (1994). The role of home practice in the thermal biofeedback treatment of migraine headache. *Journal of Consulting and Clinical Psychology, 62,* 180-184.

Gauthier, J., Doyon, J., Lacroix, R., & Drolet, M. (1983). Blood volume pulse biofeedback in the treatment of migraine headache: A controlled evaluation. *Biofeedback and Self-Regulation, 8,* 427-442.

Gauthier, J., Fradet, C., & Roberge, C. (1988). The differential effects of biofeedback in the treatment of classical and common migraine. *Headache, 28,* 39-46.

Gauthier, J.G., Ivers, H., & Carrier, S. (1996). Nonpharmacological approaches in the management of recurrent headache disorders and their comparison and combination with pharmacotherapy. *Clinical Psychology Review, 16,* 543-571.

*Gauthier, J., Lacroix, R., Cote, A., Doyon, J., & Drolet, M. (1985). Biofeedback control of migraine headaches: A comparison of two approaches. *Biofeedback and Self-Regulation, 10,* 139-159.

Gawel, M.J., Kreeft, J., Nelsen, R.F., Simard, D., & Arnott, W.S. (1992). Comparison of the efficacy and safety of Flunarizine and Propranolol in the prophylaxis of migraine. *Canadian Journal of Neurologic Science, 19,* 340-345.

Gerber, W.D., Diener, H.C., Scholz, E., & Niederberger, U. (1991). Responders and nonresponders to metoprolol, propranolol and nifedipine treatment in migraine prophylaxis: A dose-range study based on time-series analysis. *Cephalagia, 11,* 37-45.

Gerber, W.D., Schellenberg, R., Thom, M., Haufe, C., Bolsche, F., Wedekind, W., Niederberger, U., & Soyka, D. (1995). Cyclandelate versus propranolol in the prophylaxis of migraine: Double-blind placebo-controlled study. *Functional Neurology,10,* 27-35.

Germain, H., & Neuron, A. (1990) Flunarizine in headache. *Headache Quarterly, 1,* 320-323.

Ghose, K., & Niven, B. (1998). Prophylactic sodium valproate theapy in patients with drug-resistant migraine. *Methods of Experimental and Clinical Pharmacology, 20,* 353-359.

Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educational Research, 5,* 3-8.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Glasziou, P.P., & Sanders, S.L. (2002). Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine, 21,* 1503-1511.

Goadsby, P.J. (1997). How do the currently used prophylactic agents work in migraine. *Cephalagia, 17,* 85-92.

Goldstein, M., & Chen, T.C. (1982). Epidemiology of disabling headache. In A.R. Critchley, S. Friedman, F. Goring, & D.S. Sicuter (Eds.), *Advances in neurology* (pp. 377-390). New York: Raven Press.

Grotemeyer, K.H., Schlake, H.P., & Hussted, I.W. (1987). Metoprolol versus Flunarizine-A double blind cross-over study. *Cephalagia, 9*(Suppl. 6), 465-466.

Grotemeyer, K.H., Schlake, H.P., & Hussted, I.W. (1989). Normalization of platelet-reactivity under successful migraine-prophylaxis with metoprolol or flunarizine. *Cephalagia,* (Suppl. 10), 435-436.

Hansen, K., Sorensen, P.S., & Olesen, J. (1989). A controlled study of flunarizine in common migraine. *Acta Neurologica Scandinavia, 67,* 266-267.

Hart, J.D. (1984). Predicting differential response to EMG biofeedback and relaxation training: The role of cognitive structure. *Journal of Clinical Psychology, 40,* 453-457.

*Havanka-Kanniainen, H., Hokkanen, E., & Myllyla, V.V. (1988). Long acting propranolol in the prophylaxis of migraine: Comparison of the daily doses of 80 mg and 160 mg. *Headache, 28,* 607-611.

Headache Classification Committee of the International Headache Society. (1988). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalagie, 8*(Suppl. 7), 1-96.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hering, R., & Kuritzky, A. (1992). Sodium valproate in the prophylactic treatment of migraine: A double-blind study versus placebo. *Cephalagia,12,* 81-84.

Holmes, B., Brogden, R.N., Heel, R.C., Speight, T.M., & Avery, G.S. (1984). Flunarizine: A review of its pharmacodynamic and pharmacokinetic properties and therapeutic use. *Drugs, 27,* 6-44.

Holroyd, K,A. (2002). Assessment and psychological management of recurrent headache disorders. *Journal of Consulting and Clinical Psychology, 70,* 656-677.

*Holroyd, K.A., France, J.L., Cordingley, G.E., Rokicki, L., Kvaal, S., Lipchik, G., & McCool, H. (1995). Enhancing the effectiveness of relaxation-thermal biofeedback training with propranolol hydrochloride. *Journal of Consulting and Clinical Psychology, 63,* 327-330.

Holroyd, K.A., & French, D.J. (1995). Recent developments in the psychological assessment and management for recurrent headache disorders. In A.J. Goreczny (Ed.), *Handbook of health and rehabilitation psychology* (pp. 3-30). New York: Plenum.

*Holroyd, K.A., Holm, J.E., Hursey, K.G., Penzien, D.B., Cordingley, G.E., Theofanous, A.G., Richardson, S.C., & Tobin, D.L. (1988). Recurrent vascular headache: Home-based behavioral treatment versus abortive pharmacological treatment. *Journal of Consulting and Clinical Psychology, 56,* 218-223.

Holroyd, K.A., Holm, J.F., Penzien, D.B., Cordingley, G.E., Hursey, K.G., Martin, N.J., & Theofanous, A. (1989). Long-term maintenance of improvements achieved with (abortive) pharmacological and nonpharmacological treatments for migraine: Preliminary findings. *Biofeedback and Self Regulation, 14,* 301-308.

Holroyd, K.A., & Penzien, D.B. (1986). Client variables and behavioral treatment of recurrent tension headache: A meta-analytic review. *Journal of Behavioral Medicine, 9,* 515-535.

Holroyd, K.A., & Penzien, D.B. (1990). Pharmacological versus non-pharmacological prophylaxis of recurrent migraine headache: A meta-analytic review of clinical trials. *Pain, 42,* 1-13.

Holroyd, K.A., Penzien, D.B., & Cordingley, G.E. (1991). Propranolol in the management of recurrent migraine: A meta-analytic review. *Headache, 31,* 333-340.

Holroyd, K.A., Penzien, D.B., Rokicki, L.A., & Cordingley, G.E. (1992). Flunarizine vs. Propranolol: A meta-analysis of clinical trials. *Headache,32* , 256.

Howard, L., Reardon, J.P., & Tosi, D.J. (1982). Modifying migraine headache through rational stage directed hypnotherapy: A cognitive-experiential perspective. *International Journal of Clinical and Experimental Hypnosis, 30,* 257-269.

Ilaqua, G.E. (1994). Migraine headaches: Coping efficacy of guided imagery training. *Headache, 34,* 99-102.

Jadad, A.R., & McQuay, H.J. (1996). Meta-Analyses to evaluate analgesic interventions: A systematic quantitative review of their methodology. *Journal of clinical Epidemiology, 49,* 2, 235-243.

Jadad, A.R., Moore, R.A., Carol, D., Jenkinson C., Reynolds D.J., Gavaghan D.J., & McQuay, H.J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials, 17,* 1-12.

Janssen, K., & Neutgens, J. (1986). Autogenic training and progressive relaxation in the treatment of three kinds of headache. *Behaviour Research and Therapy, 24,* 199-208.

*Jensen, R., Brink, T., & Olesen, J. (1994). Sodium valproate has a prophylactic effect in migraine without aura: A triple-blind, placebo-controlled crossover study. *Neurology, 44,* 647-651.

*Johnson, R.H., Hornabrook, R.W., & Lambie, D.G. (1986). Comparison of mefamic acid and propranolol with placebo in migraine prophylaxis. *Acta Neurological Scandinavia, 73,* 490-492.

*Jurish, S.E., Blanchard, E.B., Andrasik, F., Teders, S.J., Neff, D.F., & Arena, J.G. (1983). Home- versus clinic-based treatment of vascular headache. *Journal of Consulting and Clinical Psychology, 51,* 743-751.

Kabela, E., Blanchard, E.B, Appelbaum, K.A., & Nicholson, N. (1989). Self-regulatory treatment of headaches in the elderly. *Biofeedback and Self-Regulation, 14,* 219-228.

*Kaniecki, R.G. (1997). A comparison of Divalproex with propranolol and placebo for the prophylaxis of migraine without aura. *Archives of Neurology, 54,* 1141-1145.

*Kangasniemi, P., & Hedman, C. (1984). Metoprolol and propranolol in the prophylactic treatment of classical and common migraine: A double blind study. *Cephalagia, 4,* 91-96.

*Kangasniemi, P.J., Nyrke, T., Lang, A. H., & Petersen, E. (1983). Femoxetine a new 5-HT uptake inhibitor, and propranolol in the prophylactic treatment of migraine. *Acta Neurological Scandinavia, 68*, 262-267.

Kass, B., & Nestvold, K. (1980). Propranolol (Inderal) and clonidine (Catapressan) in the prophylactic treatmentof migraine. A comparative trial. *Acta Neurologica Scandandinavia, 61,* 351-356.

Kewman, D., & Roberts, A.H. (1980). Skin temperature biofeedback and migraine headaches. *Biofeedback and Self Regulation, 5,* 327-345.

*Kinze, S., Clauss, M., Reuter, U., Wolf, T., Dreier, J.P., Einhaupl, K.M., & Arnold, G. (2001). Valproic acid is effective in migraine prophylaxis at low serum levels: A prospective open-label study. *Headache, 41*, 774-778.

Kirk-Smith, M.D., & Stretch, D.D. (2001). Evidence-based medicine and randomized double-blind clinical trials: A study of flawed implementation. *Journal of Evaluation of Clinical Practice, 7,* 119-123.

*Kjaersgard Rasmussen, M.J., Holt Larsen, B., Borg, L., Soelberg Sorensen, P., & Hansen, P.E. (1994). Tolfenamic acid versus propranolol in the prophylactic treatment of migraine. *Acta Neurological Scandinavia, 89*, 446-450.

*Klapper, J. (1997). Divalproex sodium in migraine prophylaxis: A dose-controlled study. *Cephalagia, 17,* 103-108.

Kunkel, R.S. (1987). Clinical trials in migraine: Parallel versus crossover studies. *Neuroepidemiology, 6,* 209-213.

*Kuritzky, A., & Hering, R. (1987). Prophylactic treatment of migraine with long-acting propranolol: A comparison with placebo. *Cephalagia, 7*(Suppl. 6), 457-458.

Lacroix, J.M., Clarke, M.A., Bock, L.C., Doxey, N., Wood, A., & Lavis, S. (1983). Biofeedback and relaxation in the treatment of migraine headaches: Comparative effectiveness and physiological correlates. *Journal of Neurological and Neurosurgical Psychiatry, 46*, 525-532.

*Lake, A., Raney, J., & Papsdorf, J.D. (1979). Biofeedback and rational emotive therapy in the management of migraine headache. *Journal of Applied Behavioral Analysis, 12*, 127-140.

*Lamsudin, R., & Sadjimin, T. (1993). Comparison of the efficacy between Flunarizine and Nifedipine in the prophylaxis of migraine. *Headache, 33*, 335-8.

Lance, J. W. (1993). Current concepts of migraine pathogenesis. *Neurology, 42*(Suppl. 2), S11-S15.

Lander, J. (1990). Clinical judgements in pain management. *Pain, 42*, 15-22.

Landy, S.H., & McGinnis, J. (1999). Divalproex sodium-review of prophylactic migraine efficacy, safety and dosage, with recommendations. *Tennessee Medicine, 92,* 135-136.

*Largen, J.W., Mathew, R.J., Dobbins, K., & Claghorn, J.L. (1981). Specific and nonspecific effects of skin temperature control in migraine control. *Headache,21*, 36-44.

Leahey, W.J., Neill, J.D., Varma, M.P.S., & Shanks, R.G. (1980). Comparison of the efficacy and pharmacokinetics of conventional propranolol and a long-acting preparation of propranolol. *British Journal of Clinical Pharmacology, 9*, 33-40.

Leandri, M., Parodi, C., Bacigalupo, F., & Farinini, D. (1985). Long-term study on effectiveness of flunarizine in headache prophylaxis, *Cephalagia, 5*(Suppl. 3), 556-557.

*Lenaerts, M., Bastings, E., Sianard, J., & Schoenen, J. (1996). Sodium valproate in severe migraine and tension-type headache: An open study of long-term efficacy and correlation with blood levels. *Acta Neurologica Belgium, 96*, 126-129.

Lisspers, J., & Ost, L.G. (1990). Long-term follow-up of migraine treatment: Do the effects remain up to six years? *Behavioral Research and Therapy, 28*, 313-322.

Louis, P. (1987). Effect of flunarizine on prodomes in migraine. *Cephalagia, 7*(Suppl. 6), 417.

*Louis, P. (1981). A double-blind placebo controlled prophylactic study of flunarizine (Sibelium) in migraine. *Headache, 21*, 235-239.

*Louis, P., & Spierings, E.L. (1982). Comparison of flunarizine (Sibelium) and pizotifen (Sandomigran) in migraine treatment: A double-blind study. *Cephalagia, 2,* 197-203.

*Lucking, C.H., Oestreich, W., Schmidt, R., & Soyka, D. (1988). Flunarizine vs. propranolol in the prophylaxis of migraine: Two double-blind comparative studies in more than 400 patients. *Cephalagia, 8*(Suppl. 8), 21-26.

*Ludin, H.P. (1989). Flunarizine and propranolol in the treatment of migraine. *Headache, 29*, 218-233.

Lukoff, D., Edwards, D., & Miller, M. (1998). The case study as a scientific method for researching alternative therapies. *Therileshh Medicine, 4,* 44-62.

Malone, M.D., & Strube, M.J. (1988). Meta-analysis of non-medical treatments for chronic pain. *Pain, 34*, 231-244.

Malvea, B.P., Gwon, N., & Graham, J.R. (1973). Propranolol prophylaxis of migraine. *Headache, 12,* 163-167.

Martinez-Lage, J.M. (1988). Flunarizine (Sibelium) in the prophylaxis of migraine. An open, long-term, multicenter trial. *Cephalagia, 8*(Suppl. 8),15-20.

*Mathew, N.T. (1981). Prophylaxis of migraine and mixed headache: A randomized controlled study. *Headache, 21*, 105-109.

Mathew, N.T. (2001) Pathophysiology, epidemiology, and impact of migraine. *Clinical Cornerstone, 4,* 1-17.

Mathew, N.T., & Ali, S. (1991). Valproate in the treatment of persistent chronic daily headache. An open label study. *Headache, 31,* 71-74.

*Mathew, N.T., Saper, J.R., Silberstein, S.D., Rankin, L., Markley, H.G., Solomon, S., Rapoport, A.M., Silber, C.J., & Deaton, R.L. (1995). Migraine prophylaxis with Divalproex. *Archives Neurology, 52*, 281-286.

*McGrady, A., Wauquier, McNiel, A., & Gerard, G. (1994). Effect of biofeedback-assisted relaxation on migraine headache and changes in cerebral blood flow velocity in the middle artery. *Headache, 34,* 424-428.

Mendenopoulos, G., Manafi, T., Logothetis, I., & Bostantjopoulou, S.(1985a). Flunarizine in the prevention of classical migraine: A placebo-controlled evaluation. *Cephalagia, 5*, 31-7.

*Mentenopoulos, G., Manafi, T., Logothetis, J., & Bostantzopoulou, S. (1985b). Flunarizine in the prevention of classical migraine: A placebo-controlled evaluation. *Cephalagia, 5*(Suppl. 2), 135-140.

*Mikkelsen, B., Kjaersgaard Pedersen, K., & Christiansen, L.V. (1986). Prophylactic treatment of migraine with tolfenamic acid, propranolol, and placebo. *Acta Neurological Scandinavia, 73*, 423-427.

Mitsikostas, D.D., & Polychronidis, I. (1997). Valproate versus flunarizine in migraine prophylaxis: A randomized, double-open, clinical trial. *Functional Neurology, 12,* 267-276.

Mizener, D., Thomas, M., & Billings, R.F. (1988). Cognitive changes of migraineurs receiving biofeedback training. *Headache, 28,* 339-343.

Morrill, B., & Blanchard, E.B. (1989). Two studies of the potential mechanisms of action in the thermal biofeedback treatment of vascular headache. *Headache, 29,* 169-176.

Mullerners, W.M., Whitmarsh, T.E., & Steiner, T.J. (1998). Noncompliance may render migraine prophylaxis useless, but once-daily regimes are better. *Cephalagia, 18,* 52-56.

*Mullinix, J., Norton, B., Hack, S., & Fishman, M. (1978). Skin temperature biofeedback and migraine. *Headache, 17,* 242-244.

*Nadelmann, J.W., Phil, M., Stevens, J., & Saper, J.R. (1986). Propranolol ® in the prophylaxis of migraine. *Headache, 26,* 175-182.

Nappi, G., Sandrini, G., Savoini, G., Cavallini, A., de Rysky, C., & Micieli, G. (1987). Comparative efficacy of cyclandelate versus flunarizine in the prophylactic treatment of migraine. *Drugs, 33*(Suppl. 2), 103-109.

Nicholson, N.L., & Blanchard, E.B. (1993). A controlled evaluation of behavioral treatment of chronic headache in the elderly. *Behavior Therapy, 24,* 395-408.

*Nuti, A., Lucetti, C., Pavese, N., Dell'Agnello, G., Rossi, G., & Bonuccelli, U. (1996). Long-term follow-up after flunarizine or nimodipine discontinuation in migraine patients. *Cephalagia,16,* 337-340

Olerud, B., Gustavsson, C.L., & Furberg, B. (1986). Nadolol and propranolol in migraine management. *Headache, 26,* 490-493.

Olson, R.P. (1988). A long-term, single group follow-up study of biofeedback therapy with chronic medical and psychiatric patients. *Biofeedback and Self-Regulation,13,* 331-346.

Olsson, J.E., Behring, H.C., Forssman, B., Hedman, C., Hedman, G., Johansson, F., Kinnman, J., Palhagen, S.E., Samuelsson, M., & Strandman, E. (1984). Metoprolol and propranolol in migraine prophylaxis: A double-blind multicentre study. *Acta Neurologica Scandinavia, 70,* 160-1688.

Onghena, P., & Van Houdenhove, B. (1992). Anti-depressant induced analgesia in chronic nonmalignant pain: A meta-analysis of 39 placebo controlled studies. *Pain, 42,* 205-219.

Osterhaus, J.T., & Townsend, R.J. (1991). The quuality of life of migraineurs: A cross-sectional profile. *Cephalagia, 1*(Suppl. 11), 103-104.

Palferman, T.G., Gibberd, F.B., & Simmonds, J.P. (1983). Prophylactic propranolol in the treatment of headache. *The British Journal of Clinical Practice, 37*, 28-31.

Pascual, J., Polo, J.M., & Berciano, J. (1989). The dose of propranolol for migraine prophylaxis. Efficacy of low doses. *Cephalagia, 9,* 287-291.

Patterson, C.H. (1967). Divergence and convergence in psychotherapy. *American Journal of Psychotherapy, XXI,* 4-17.

Pfaffenrath, V., Oestreich, W., & Haase, W. (1990). Flunarizine (10 and 20 mg) i.v. versus placebo in the treatment of acute migraine attacks: A multicenter double-blind study. *Cephalagia, 10,* 77-81.

*Pini, L. A., Ferrari, A., Guidetti, G., Galetti, G., & Sternieri, E. (1985). Influence of flunarizine on the altered electronystagmographic (ENG) recordings in migraine. *Cephalagia, 5*(Suppl. 2), 173-175.

*Pita, E., Higueras, A., Bolanos, J., Perez, N., & Mundo, A. (1997). Propranolol and migraine. A clinical trial. *Archives Pharmacology and Toxicology, 3*, 273-278.

*Pradalier, A., Serratrice, G., Collard, M., Hirsch, E., Feve, J., Masson, M., Dry, J., Koulikosvsky, G., Nguyen, G., Schbath, J., & Carpentier, M.C. (1989). Long-acting propranolol in migraine prophylaxis: Results of a double-blind, placebo-controlled study. *Cephalagia, 9*, 247-253.

Rabkin, R., Stables, D.P., Levin, N.W., & Suzman, M.M. (1966). The prophylactic value of propranolol in angina pectoris. *American Journal of Cardiology, 18,* 370-383.

*Rao, B.S., Das, D.G., Taraknath, V.R., & Sarma, Y. (2000). A double-blind controlled study of propranolol and cyproheptadine in migraine prophylaxis. *Microencapulation Neurological Indian, 48,* 223-226.

*Rascol, A., Montastruc, J.L., & Rascol, O. (1985). Flunarizine vs pizotifen: A double-blind study in the prophylaxis of migraine. *Headache, 26*, 83-85.

Raskin, N.H. (1988) *Headache*. New York: Churchill Livingstone.

Raskin, N.H., & Schwartz, R.K. (1980). Interval therapy of migraine: Long-term results. *Headache, 20,* 336-340.

Rasmussen, B., Jensen, R., & Olesen, J. (1991). Questionnaire versus clinical interview in the diagnosis of headache. *Headache, 31,* 290-295.

Reading, C. (1984). Psychophysiological reactivity in migraine following biofeedback. *Headache, 24,* 70-74.

Reich, B.A. (1989). Noninvasive treatment of vascular and muscle contraction headache: A comparative longitudinal clinical study. *Headache, 29,* 34-41.

Richardson, G.M., & McGrath, P.J. (1989). Cognitive-behavioral therapy for migraine headaches: A minimal-therapist-contact approach versus a clinic-based approach. *Headache, 29,* 352-357.

*Rosen, J. (1983). Observations on the efficacy of propranolol for the prophylaxis of migraine. *Annuals of Neurology, 13,* 92-93.

*Rothrock, J.F., Kelly, N.M., Brody, M.L., & Golbeck, A. (1994). A differential response to treatment with Divalproex sodium in patients with intractable headache. *Cephalagia, 14,* 241-244.

Rothrock, J.F., & Mendizabal, J.E. (2000). An analysis of the "carry-over effect" following successful short-term treatment of transformed migraine with divalproex sodium. *Headache, 40,* 17-19.

Rudner, L.M., Glass, C., Evart, D.L., & Emery, T.J. (2000). Meta-Stat ® [Computer statistical software]. Retrieved November 4, 2002, from http://www.ericae.net/meta/preface.htm

Sappington, J.T., & Fiorito, E.M. (1985). Thermal feedback in Raynaud's phenomenon secondary to systematic lupus erytheratosus: Long-term remission of target symptoms. *Biofeedback and Self Regulation, 10,* 335-341.

*Sargent, J., Solbach, P., Damasio, H., Baumel, B., Corbett, J., Eisner, L., Jessen, B., Kudrow, L., Nathew, N., Medina, J., Saper, J., Vijayan, N., Watson, C., & Alger, J. (1985). A comparison of Naproxen Sodium to Propranolol Hydrochloride and a placebo control for the prophylaxis of migraine headache. *Headache, 25,* 320-324.

Sargent, J.D., Walters, E.D., & Green, E.E. (1973). Psychosomatic self-regulation of migraine headaches. *Seminars in Psychiatry, 5,* 415-428.

Schellenberg, R., Milch, W., Schwarz, A., & Schoberg, F. (1994). Quanititative Eeg and BPRS data followng Haldol-Decanoate (R) administration in schizophrenia. *International Clinical Psychopharmacology, 9*, 17-24.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47*, 1173-1181.

Schmidt, R., & Oestreich, W. (1991) Flunarizine in migraine prophylaxis: The clinical experience. *Journal of Cardiovascular Pharmacology, 18*(Suppl. 8), S21-26.

Scholz, E., Gerber, W.D., Bille, A., Niederberger, U., & Fahrner, I. (1987). Plasma levels of metoprolol and propranolol in responders and non-responders to prophylactic treatment of migraine. *Cephalagia, 7*(Suppl. 6), 412-413.

Senger, H.L. (1987). The placebo effect of psychotherapy: A moose in the rabbit stew. *American Journal of Psychotherapy, XLI,* 68-81.

Setter, F., & Kupper, S. (2002). Autogenic training: A meta-analysis of clinical outcome studies. *Applied Psychophysiology and Biofeedback, 27*, 45-98.

Shapiro, A.K., & Morris, L.A. (1978). The placebo effect in medical and psychological therapies. In S.L. Bergin & A.E. Garfield, *Handbook of psychotherapy and behavior change, an empirical analysis* (2nd ed., pp. 369-411). New York: Wiley.

Sheldrake, R. (1998). Experimenter effects in scientific research: How widely are they neglected? *Journal of Scientific Exploration, 12*, 73-78.

Shelton, C.E., & Connelly, J.F. (1996). Valproic acid: A migraine prophylaxis alternative. *Annals of Pharmacotherapy, 30,* 865-866.

Shimell, C.J., Fritz, V.U., & Levien, S.L. (1990) A comparative trial of flunarizine and propranolol in the prevention of migraine. *South African Medical Journal, 77,* 75-77.

Silberstein, S.D., & Collins, S.D. (1999). Safety of divalproex sodium in migraine prophylaxis: An open-label, long-term study. Long-term safety of Depakote in headache prophylaxis study group. *Headache, 39,* 633-643.

Silberstein, S.D., & Wilmore, L.J. (1996). Divalproex sodium: Migraine treatment and monitoring. *Headache, 36,* 239-242.

Slavin, R.E. (1984). Meta-analysis in education: How has it been used? *Educational Researcher, 13,* 6-27.

Smith, W.B. (1987). Biofeedback and relaxation training: The effect on headache and associated symptoms. *Headache, 27*, 511-514.

Solomon, G.D. (1986). Verapamil and Propranolol in migraine prophylaxis: A double blind, crossover study. *Headache, 26,* 325.

Solomon, G.D. (1995). The pharmacology of medications used in treating headache. *Seminal Pediatric Neurology, 2,* 165-177.

Sorbi, M., & Tellegen, B. (1984).Ways of appraisal and coping in migraine provoking situations. *Gedrad: Tijdschrift voor Psychologie, 12,* 5-15.

Sorbi, M., & Tellegen, B. (1986). Differential effects of training in relaxation and stress-coping in patients with migraine. *Headache, 26,* 473-481.

Sorbi, M., Tellegen, B., & DuLong, A. (1989). Long-term effects of training in relaxation and stress-coping in patients with migraine: A 3-year follow-up. *Headache, 29,* 111-121.

Sorensen, K.V. (1988). Valproate: A new drug in migraine prophylaxis. *Acta Neurologica Scandinavia, 78,* 346-348.

*Sorensen, P.S., Hansen, K., & Olesen, J. (1986). A placebo-controlled, double-blind, cross-over trial of flunarizine in common migraine. *Cephalagia, 6,* 7-14.

Sorensen, P.S., & The Danish Migraine Study Group. (1989). Prophylactic effect of flunarizine vs. metoprolol in migraine. *Cephalagia, 9*(Suppl. 10), 103-109.

Sovak, M., Kunzel, M., Sternbach, R.A., & Dalessio, D.J. (1981). Mechanism of the biofeedback therapy of migraine: Volitional manipulation of the psychophysiological background. *Headache, 21,* 89-92.

Soyka, D., & Oestreich, W. (1987) Therapeutic effectiveness of flunarizine and propranolol in the interval therapy of migraine. *Cephalagia, 7*(Suppl. 6), 467-468.

Soyka, D., Taneri, Z., Oestreich, W., & Schmidt, R. (1988). Flunarizine i.v. in the acute treatment of the migraine attack. A double-blind placebo-controlled study. *Cephalagia, 8*(Suppl. 8), 35-40.

Spierings, E.L., & Messinger, H.B. (1988). Flunarizine vs. pizotifen in migraine prophylaxis: A review of comparative studies, *Cephalagia, 8*(Suppl. 8), 27-30.

*Standness, B. (1982). The prophylactic effect of timolol versus propranolol and placebo in common migraine: Beta blockers in migraine. *Cephalagia, 2*, 165-170.

Steardo, L., Marano, E., Barone, P., Denman, D.W., Monteleone, P., & Cardone, G. (1986). Prophylaxis of migraine attacks with a calcium-channel blocker: Flunarizine versus methysergide. *Journal of Clinical Pharmacology, 26,* 524-528.

*Stensrud, P., & Sjaastad, O. (1976). Short-term clinical trial of propranolol in racemic form (inderal), d-propranolol and placebo in migraine. *Acta Neurological Scandinavia, 53*, 229-232.

Stensrud, P., & Sjaastad, O. (1980a). Comparative trial of Tenormin (atenolol) and Inderal (propranolol) in migraine. *Headache, 20,* 204-207.

Stensrud, P., & Sjaastad, O. (1980b). Comparative trial of Tenormin (atenolol) and Inderal (propranolol) in migraine. *Journal of Medical Science, 31,* 37-40.

Stewart, W.F., Shechter, A., & Rasmussen, B.K. (1994). Migraine prevalence. A review of population-based studies. *Neurology, 44*(6)(Suppl. 4), S17-23.

Sudilovsky, A., Elkind, A.H., Ryan, R.E. Sr., Saper, J.R., Stern, M.A., & Meyer, J.H. (1987). Comparative efficacy of nadolol and propranolol in the management of migraine. *Headache, 27,* 421-426.

Taylor, S.E. (1991). *Health psychology* (2nd ed.). New York: McGraw-Hill.

*Tfelt-Hansen, P., Standnes, B., Kangasneimi, P., Hakkarainen, H., & Olesen, J. (1984). Timolol vs. propranolol vs. placebo in common migraine prophylaxis: A double-blind multicenter trail. *Acta Neurologica Scandinavia, 69*, 1-8.

Thomas, R.C. (1989). Proton channels in snail neurones. Does calcium entry mimic the effects of proton influx? *Annals of New York Academy Science, 574*, 287-293.

*Thomas, M., Behari, M., & Ahuja, G. K. (1991). Flunarizine in migraine prophylaxis: An Indian trial. *Headache, 31*, 613-615.

Todd, P.A., & Benfield P. (1989). Flunarizine: A reappraisal of its pharmacological properties and therapeutic use in neurological disorder. *Drugs, 38*, 481-499.

Turner, J.A., Jensen, M.P., Warms, C.A., & Cardenas, D.D. (2002). Blinding effectiveness and association of pretreatment expectations with pain improvement in a double-blind randomized controlled trial. *Pain, 99*, 91-99.

Turin, A., & Johnson, W.G. (1975). Biofeedback treatment of migraine headache: A systematic case study. *Behavior Therapy, 6,* 394-397.

Unruh, A.M. (1996). Gender variations in clinical pain experience. *Pain, 65*, 123-167.

Ward, T.N. (2000) Providing relief from headache pain. Current options for acute and prophylactic therapy. *Postgraduate Medicine, 108*, 121-128.

Wauquier, A., Ashton, D., & Marrannes, R. (1985). The effects of flunarizine in experimental models related to the pathogenesis of migraine. *Cephalagia, 5* (Suppl. 2), 119-123.

*Wauquier, A., McGrady, A., Aloe, L., Klauser, T., & Collins, B. (1995). Changes in cerebral block flow velocity associated with biofeedback-assisted relaxation treatment of migraine headaches are specific for the middle cerebral artery. *Headache, 35*, 358-362.

*Weber, R.B., & Reinmuth, O.M. (1972). The treatment of migraine with propranolol. *Neurology, 22*, 366-369.

Weillner, C., May, A., Limmroth, V., Jupter, M., Kaube, H., & Schayak, R.V. (1995). Brain stem activation in spontaneous human migraine attacks. *Nature Medicine, I,* 658-660.

Werbach, M.R., & Sandweiss, J.H. (1978). Peripheral temperatures of migraineurs undergoing relaxation training. *Headache, 18,* 211-214.

*Wilderoe, T., & Vigander, T. (1974). Propranolol in the treatment of migraine. *British Medical Journal, 2*, 669-701.

Wober, C., Wober-Bingol, C., Koch, G., & Wessely, P. (1991). Long-term results of migraine prophylaxis with flunarizine and beta-blockers. *Cephalagia, 11*, 25-26.

*Ziegler, D.K., Hurwitz, A., Preskorn, S., Hassanein, R., & Seim, J. (1993). Propranolol and Amitryptyline in prophylaxis of migraine. *Archives of Neurology, 50*, 825-830.

APPENDICES

Appendix A: Pharmacological Agents' Names

Table A1

Pharmacological Agents' Names

Category	Brand name	Generic name	Chemical name
Beta-adrenergic blocker	Inderal®	Propranonol	*(Isopropylamino)-3-(1-naphthyloxy)-2-propanol hydrochloride.
Calcium channel blocker	Sibelium ®	Flunarizine	*(4-fluorophenyl)-Methyl-4-(3-phenyl-2-propenyl) *piperazine-dihydrochloride.
Anti-convulsant	Depakote ®	Divalproex Sodium	*sodium hydrogen (2-propylpentanoate). *valproic acid *sodium valproate

Appendix B: Summary of Reviews

Table B1

Summary of Reviews

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Blanchard, Andrasik, Ahles, Teders, & O'Keefe (1980)	migraine	biofeedback, relaxation, bio/relax .	47	yes	pre-post change scores	gender, age, quality, duration of treatment.	All three treatments show similar effectiveness rates. Outcome does not vary by gender or age Inconclusive on duration of treatment.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Bogaards & ter Kuile (1994)	tension headache	EMG Biofeedback, Relaxation, Bio/relax, Cognitive Therapy. Pharmaco-logical (amitriptyline, diazepam, ibuprofen, asprin, aceta-minophen, naproxen-sodium, clomipramine, doxepin, trzanidine) Placebo, Control.	78	yes	pre-post treatment scores	treatment setting, duration of treatment, therapist training, age, duration of headache, gender, method of subject recruitment, number of subjects, drop outs, outcome measure, diagnostic criteria, year of publication, internal validity.	All non-pharmacological treatments are superior to pharmacological treatments for tension headache (when pharmacological treatments are pooled). Outcomes vary by type of outcome measure used (headache diary versus other methods. Studies that examine short duration treatments result in higher percentages of improvement. Age is related to poor outcome. Gender and method of recruitment are not related to outcome.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Compas, Haag, Keefe, Leintenberg, & Williams (1998)	migraine	Cognitive therapy, Cog-beh therapy, biofeedback, relaxation.	21	no	pre- post improvement	none	All biofeedback and relaxation treatments have similar effect rates. There is limited evidence for Cognitive therapies, but what is available appears promising.
Duke University & The Center for Clinical Health Policy Research (1999a)	migraine	Relaxation therapy, thermal biofeedback, EMG biofeedback, cognitive-behavioral therapy, acupuncture, TENS, spinal manipul-ation, hyperbaric oxygen.	29	yes	Standard-ed effect sizes percent improved	none	Thermal biofeedback plus relaxation therapy, relaxation training, and EMG biofeedback are al modestly effective in treating migraine.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Duke University & The Center for Clinical Health Policy Research (1999a)	migraine	(Category-see detailed list at the end of the table) alpha-2-agonsits, anitconvul-sants, anti-depressants, beta-blockers, calcium antagonists, ergots, NSAIDS, hormonal treatments, and seritonin-ergic drugs.	258	yes	Standardized effect sizes, Odds ratios	none	Propranolol, timinolol, divalproex sodium, amitriptyline, flunarizine, and pizotifen all have effect sizes between .52 and .93. Participant drop outs are a major issue in pharmacological research.
Flor, Fydrich, & Turk (1992)	back pain	no headache treatments examined	65	yes	Standardized effect sizes	age, gender, pain duration, % working, % married, years in school, litigation, compensation, surgery, long/short term.	Outcome varies by age, marital status, education level, SES, employment stats, compensation and medication use.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Holroyd & Penzien (1985)	tension and client variables	EMG Biofeedback, Relaxation, Combo treatment, no treatment,	37	yes	Pre-post change scores (only from studies using headache diaries)	# of subjects, dropout, gender, age, source of client, duration of treatment, transfer training, assignment to groups, internal validity, diagnostic criteria.	Outcome does not vary by research design. Outcome varies by age, year of publication, gender, and dropout rate. All treatments have similar effectiveness rates.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Holroyd & Penzien (1990)	migraine	Propranonol, bio/relax, no treatment, placebo.	60	yes	headache index pre-post scores	Outcome measure.	43% improvement for both active treatments. Treatment effects vary by outcome measure.
Holroyd & Penzien (1990)	migraine	Propranonol, relaxation/ biofeedback, placebo.	73	yes	headache index pre-post improvement score	number of subjects, gender, age, outcome measure, dropout rate, migraine diagnosis, Propranonol dose, dropout due to side effects, blind to condition, # of bio sessions, transfer training.	Similar improvement rates for biofeedback/relax ation training and Propranonol. Results vary by type of outcome measure (daily recording versus global estimates).

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Holroyd, Penzien, & Cordingley (1991)	migraine	Propranonol.	53	yes	pre-post scores, placebo-treatment change score	# of subjects, age, migraine diagnosis, chronicity, dropout rate, drop out due to side effects, dose, blind to condition.	Conservative estimates of Propranonol treatment effects indicate a 44% reduction in frequency, duration, and intensity of migraine headache. Dose issues do not effect outcome. Type of outcome measure do effect treatment outcome (headache diary versus global measurements)
Lander (1980)	general pain	No headache treatments considered.	42	no	none	patient variables, differences in sample, selection, gender, SES, age, duration of pain	Research on pain management is poor due to methodological flaws and a failure to include the referenced variables.

(*table continues*)

Authors (date)	Type of headache	Types of treatment	Number of studies analyzed	Quantitative (yes/no)	Type of change score used	Other variables considered	Major findings
Onghena & Van Houdenhove (1992)	tension headache, migraine, chronic pain.	amitriptyline, phenelzine, imipramine, doxepin, dibenzepine, clomipramine, femoxetine, dothiepin, mianserin, trazodone,	39	yes	Standardized effect size	dosage level, depression diagnosis, # of patients, gender, mean age, mean duration of pain, quality, use of analgesics.	Antidepressant medications induce an analgesic effect. The effect decreases pain in all of the pain conditions tested.
Malone & Strube (1988)	non-Chronic pain	atogenic biofeedback, cognitive, hypnosis, no treatment, operant, pill placebo, relaxation, wait list.	109	yes	standardized effect size, percentage improved (pre-post)	type of pain (back, neck, cancer, dental, iatrogenic, join, migraine, mixed group, mixed headache, tension headache , other).	Outcomes vary by how researchers calculate change (effect size or % improved). Effect size is more sensitive. Treatments are uniform in outcome across pain categories.

(*table continues*)

Treatments examined in Duke University and the Center for Clinical Health Policy Research.

Alpha-2-agonistsis

clonidine, guanfacine.

Anticonvulsants

divalproex sodium, carbamazepine, clonazepam, gabapentin.

Antidepressants

amitriptyline, clomipramine, femoxetine, fluoxetine, fluvoxamine, mianserin, opipramol.

Beta-blockers

propranolol, metoprolol acebutolol, alprenolol, atenolol, bisoprolol, nadolol, oxprenolol, pindolol, practolol, timolol.

Calcium antagonists

cyclandelate, flunarizine, nicardipine, nifedipine, nimodipine, verapamil.

Ergots

dihydroergotamine, dihydroergokrptine, egrotamine, Cafergot

Methysergide; nonsteroidal anti-inflammatory drugs (NSAIDS)

aspirin, fenoprefen, flurbiprofen, indobufen, idomethacin, ketoprefen, lornoxicam, mefamic acid, naproxen, naproexn sodium, tolfenamic acid.

Other Serotoninergic Drugs

pizotifen, lisuride, oxitriptan, iprazochrome, tropisetron.

Other Treatments

hormonal preparations and feverfew.

Appendix C: Excluded Articles

Table C1

Propranolol Excluded Articles

Authors (date)	Reason for exclusion
Al Qassab & Findley, 1992	Insufficient data
Baldrati et al., 1983	Insufficient data
Behan & Reid, 1980	Insufficient data
Bordini, Arruda, Ciciarelli, & Speciali, 1997	Insufficient data
Carroll, Reidy, Savundra, Cleave, & McAinsh, 1990	Insufficient data
Cortelli et al., 1985	Insufficient data
Diamond, Kudrow, Stevens, & Shapiro, 1982	Insufficient data
Diener et al., 2002	Insufficient data
Gawel, Kreeft, Nelson, Simard, & Arnott, 1992	Insufficient data
Gerber, Diener, Scholz, & Niederberger, 1991	Insufficient data
Gerber et al., 1995	Insufficient data
Holroyd, Penzien, Rokicki, & Cordingley, 1992	Brief report-insufficient data
Kass & Nestvold, 1980	Insufficient data
Leahey, Neill, Varma & Shanks, 1980	Results on blood pressure
Malvea, Gwon, & Graham, 1973	Insufficient data
Olerud, Gustavsson, & Furberg, 1986	Insufficient data-no control group
Olsson et al., 1984	Insufficient data-no control group
Palferman, Gibberd, & Simmonds, 1983	Insufficient data

(*table continues*)

Authors (date)	Reason for exclusion
Pascaul, Polo, & Berciano, 1989	Insufficient data
Raskin & Schwartz, 1980	Retrospective study
Scholz, Gerber, Billie, Niederberger, & Fahrner, 1987	Report results for plasma levels only
Shimell, Fritz, & Levien, 1989	Insufficient data
Solomon, 1986	Abstract-unusable statistics
Stensrud & Sjaastad, 1980a	Insufficient data
Stensrud & Sjaastad, 1980b	Duplicate of above
Sudilovsky et al., 1987	Cost effectiveness study
Wobeer, Wober-Biingoi, Koch, & Wessely, 1991	Long term results after discontinuation of treatment.

Table C2

Flunarizine Excluded Articles

Authors (date)	Reason for exclusion
Amery, 1983	Reviews others data
Amery et al., 1981	Reviews pharmacology
Angoli et al., 1991	Comparison-provided insufficient data
Andersson, 1985	Abstract, insufficient data
Baker, 1987	Abstract, insufficient data
Bassi, Brunati, Rapuzzi, Alberti, & Mangoni, 1992	Insufficient data
Bonuso et al., 1986	Migraines induced
Centonze et al., 1983	Insufficient data
Centonze, Magrone, & Vino, 1990	Insufficient data
Cerbo et al., 1986	Insufficient data
D'Amato, D'Amato, Alfano, Giordano, & Marmo, 1990	Insufficient data
Diamond & Schenbaum, 1983	Insufficient data
Germain & Neuron, 1990	Insufficient data
Grotemeyer, Schlake, & Husstedt,.1987	Insufficient data
Grotemeyer, Schlake, & Husstedt, 1989	Insufficient data
Hansen, Sorensen, & Olesen, 1989	Insufficient data
Holmes, Brogeden, Heel, Speight, & Avery, 1984	Review
Leandri, Parodi, Bacigalupo, & Farinini, 1985	Data artificially dichotomized
Louis, 1987	Insufficient data
Louis and Spierings, 1982	Insufficient data

(table continues)

Authors (date)	Reason for exclusion
Martinez-Lage, 1988	Included large numbers of dhildren
Mendenopoulos, Manafi, Logothetis, & Bostantjopoulou, 1985	Duplicate of Mentenopoulos, Manafi, Logothetis, & Bostantjopoulou, 1985
Nappi et al., 1987	Comparison-provides insufficient data
Pfaffenrath, Oestreich, & Haase, 1990	Focus on acute treatment
Schmidt & Oestreich, 1991	Clinical case review
Sorenson and the Danish Migraine Study Group, 1989	Comparison-provides insufficient data
Soyka & Oestreich, 1987	Insufficient data
Soyka, Taneri, Oestreich, & Schmidt, 1988	Focus on acute treatment
Spierings & Messinger, 1988	Reviews others data
Sterdo et al., 1986	Comparison-provides insufficient data
Wauquier, Ashton, & Marranes, 1985	Focused on animal models

Table C3

Excluded Divalpoex Sodium Articles

Authors (dates)	Reason for exclusion
Hering & Kuritzky, 1992	Insufficient data
Klapper, 1997	Insufficient data
Mathew & Ali, 1991	Insufficient data
Mitsikostas & Polychronidis, 1997	Insufficient data
Silberstein & Collins, 1999	Insufficient data
Sorensen, 1988	Insufficient data

Table C4

Nonpharmacological Excluded Studies

Authors (dates)	Reason for exclusion
Alder & Alder, 1976	Inappropriate design
Anderson, Basker, & Dalton, 1975	Treatment not included
Andrasik, Blanchard, et al., 1984	Insufficient data
Andrasik, Pallmeyer, Blanchard, & Attanasio, 1984	Not outcome based
Andreychuk & Skriver, 1975	Insufficient data
Billings, Thomas, Rapp, Reyes, & Leith, 1984	Retrospective
Blanchard, 1987	Chronic headache
Blanchard, Andrasik, Evans, et al., 1985c	Insufficient data
Blanchard, Andrasik, Neff, Arena, et al., 1982	Inappropriate design
Blanchard, Andrasik, Neff, & Appelbaum, 1985	Inappropriate design
Blanchard, Andrasik, Neff, Teders, et al., 1982	Inappropriate design
Blanchard et al., 1987	Follow-up
Blanchard, Jaccard, Andrasik, Guarnieri, & Jurish, 1985	Focus on expenses
Brown, 1984	Treatment not included
Cohen, McArthur, & Rickles, 1980	Insufficient data
Daly, Donn, Galliher, & Zimmerman, 983	Insufficient data
Diamond & Montrose, 1983	Retrospective
Fahrion, 1977	Insufficient data
Feuerstein & Adams, 1977	Single subject
French, Gauthier, Roberge, Bouchard, & Nowen, 1997	Insufficient data

(*table continues*)

Authors (dates)	Reason for exclusion
Friedman & Taub, 1984	Insufficient data
Ford, Stroebel, Strong, & Szarek, 1983	Insufficient data
Gainer, 1978	Single subject
Gallagher & Warner, 1984	Insufficient data
Gamble & Elder, 1983	Insufficient data
Gauthier, Bois, Allaire, & Drolet, 1981	Insufficient data
Gauthier & Carrier, 1991	Follow-up
Gauthier, Doyon, Lacroix, & Drolet, 1983	Treatment not included
Gauthier, Fradet, & Roberge, 1988	Insufficient data
Gauthier, Ivers, & Carrier, 1996	Review
Hart, 1984	Insufficient data
Holroyd et al., 1989	Used only successfully treated
Howard, Reardon, & Tosi, 1982	Single subject
Ilacqua, 1994	Insufficient data- no frequency data
Janssen & Neutgens, 1986	Insufficient data
Kabela, Blanchard, Appelbaum, & Nicholson (1989)	Too few migraine subjects
Kewman & Roberts, 1980)	Insufficient data
Lacroix et al., 1983	Insufficient data
Lisspers & Ost, 1990	Follow-up
Mizener, Thomas, & Billings, 1988	Insufficient data- no frequency data
Morrill & Blanchard, 1989	Insufficient data- focus on temperature
Nicholson & Blanchard, 1993	Too few migraine subjects

(*table continues*)

Authors (dates)	Reason for exclusion
Olson, 1988	Follow-up on mixed headache
Reading, 1984	Insufficient data
Reich, 1989	Insufficient data
Sargent, Walters, & Green, 1973	Insufficient data
Smith, 1987	Insufficient data
Sorbi & Tellegen, 1984	Insufficient data
Sorbi & Tellegen, 1986	Treatment not included
Sovak, Kunzel, Sternbach, & Dalessio, 1981	Insufficient data
Turin & Johnson, 1975	Too few subjects
Werbach & Sandweiss, 1978	Insufficient data

Appendix D: Data Collection Materials

Table D1

Coding Sheet

Study ID #
Headache type 1=migraine 2=mixed 3=vascular
treatment type 1=pharm 2=nonpharm
subgroup 1=Propranonol 2=flunarizine 3=Divalproex Sodium 4=TBF 5=Relaxation 6=EMG 7=Cog therapies
% sample female
Type of headache record (1=daily diary 2=client estimation 3=researcher estimate).
Treatment length in weeks
Comparitive group 1=placebo 2=control 3=baseline-control 4=baseline comparison
Study type 1=parallel 2=crossover
Year of study

(*table continues*)

Study ID #
# of months between treatment and follow-up
# of days drug wash out period
# of sessions 2
Total mortality rate
Data type 1=continuous 2=binary
Stats-mean
sd, s, ci
t, F, Z, r
pre-post
within

Table D2

Data Page

$$Sp = \sqrt{\frac{(n1-1)s1^2 + (n2-1)s2^2}{n1+n2-2}}$$

n1= n2=

s1(sqr)= s2(sqr)=

Sp=_____ =

	Experimental	Control
N		
Mean		
F		
T		
P(exact)		
paired T		
pre		
post		
Binary		
%Success		

Appendix E: Quality of Study Rating Sheets

	Question	Response	Score
1	Was the study described as **randomized** (this includes the use of words such as randomly, random and randomization)	yes no	1 0
1a	If the method of generating the sequences of randomization was described, was it **adequate** (table of random numbers, computer-generated, coin tossing etc.) or **Inadequate** (allocated alternately, according to date of birth, hospital number, etc)	Not described Adequate Inadequate	0 1 -1
2	Was the study described as **double blind**?	Yes No	1 0
2a	If the method of blinding was described, was it **adequate** (identical placebo, active placebo, dummy etc) or **inadequate** (comparison of tablet vs. injection with no double dummy)	Not Described Adequate Inadequate	0 1 -1
3	Was there a description of withdrawals and dropouts	Yes No	1 0

Score=

Note. For sections 1a and 2a a study automatically receives a zero in these sections if the previous section (1or 2) is a zero. The scores of (-1) are only used when the study scored a (1) on the previous section. For example, a study that specifically states that it is randomized but then states it used alternate allocation would receive a score of "1" in section (1) and a score of a "-1" section (1a). Therefore, no study can receive a score of less than zero. Alternatively, a study that specifically states that it is a double blind study but does not describe how double blinding was achieved would receive a score of "1" in section (2) and a score of " 0" in section (2a).

Figure E1. Quality of study instrument (Jadad et al., 1996).

NECESSARY CONDITIONS

	1 point	**0 points**
Measurement of Headaches	____Journal/Likert Scale	_____ Self Report
Type of Control	____ Placebo/Comparison	_____ Baseline

OTHER CONDITIONS

	Yes=1	No= 0
Cross-Over Design	____	____
Run-in/Wash Out Period	____	____
Baseline/follow-up		
Multiple Setting	____	____
Compared to another Active tx.	____	____
Other Meds Controlled	____	____
Compliance Measured/ Homepractice	____	____

3=Great= All Necessary conditions and 3 other
2=Good = at least 1 of Necessary conditions and 3 other or 2 Necessary and 2 other
1=Okay = no Necessary conditions or 1 Necessary condition and less than 3 other
Nec. Cond = + =

Figure E2. Headache literature quality rating.

Appendix F: Formulas Used in Calculations of Effect Size

All calculations were performed with the aide of Meta-Stat ® a statistical software program developed by Rudner, Glass, Evart, and Emery (2000).

Calculating an unbiased standardized mean difference (*d*)

Equation 1

$$g = \overline{X}\,\text{exp} - \overline{X}cont / SD_{pooled}$$

Equation 2

$$d = g(1-(3/4m-1))$$

Example 1

So, if an article presents and experimental mean of 3.69 (N = 80) and a control mean of 2.54 (N = 80). A pooled standard deviation of 3.65. Results are as follows,

g = 3.69 - 2.54/3.65 = .31

d = .31 (1-(3/4(158)-1)

d = .31 (1-(3/631)

d = .308

Calculating an effect size from Parametric Gain Scores

Equation 3

$$d = \frac{\overline{X}_{epost} - \overline{X}_{epre}}{\dot{O}_{\exp re}} - \frac{\overline{X}_{contpost} - \overline{X}_{contpre}}{\dot{O}_{contpre}}$$

Example 2

So if the article presents experimental treatment group data as

Pre Mean = 4.12, SD = 3.7, Post Mean 1.57, SD = 1.35

and control group data as Pre Mean = 3.4, SD = 2.98, Post Mean 2.72, SD = 1.78

The calculation would be as follows

d = (4.12-1.57)/3.7 - (3.4-2.72)/3.4

d = 2.55/3.7 - .68/3.4

d = .689 - .2

d = .489

Calculating an effect size as a paired comparison t-test

Equation 4

$$d = \overline{D}\frac{\sqrt{N}}{t}$$ Where $\overline{D}$ is the mean difference between the observed pairs.

Example 3

So, for an article that presents a paired t score of 2.114, a mean difference score of .2617, with 30 paired observations, the calculation would be

$$d = .2617\sqrt{30} / 2.114$$

d = .678

Calculating an effect size from r

Equation 5

$$d = \sqrt{\frac{4r^2}{1 - r^2}}$$

See formula use in calculations below for p value and F value conversions.

The results from these formulas are plugged into the formula to calculate an effect size for r.

Calculating an r from an F-statistic

Equation 6

$$r = \sqrt{\frac{F}{F - df}}$$ Where df is the error term degree of freedom and there is only 1

degree of freedom in the interaction term.

Example 4

So, if the F value is given as 3.175 with 18 degrees of freedom in the denominator, the calculation would be as follows

$$r = \sqrt{\frac{3.175}{3.175 - 18}}$$

r = .462

Using the formula for an effect size conversion for r

$$d = \sqrt{\frac{4(.462)^2}{1 - (.462)^2}}$$

d = 1.04

Calculating an r from an exact p value

Equation 7

$$a = \sqrt{\ln(1 / p^2)}$$

Equation 8

$$z = a\left[\frac{(2.515517)(.802853a)(.010328a^2)}{1 - (1.432788a)(.189269a^2)(.001308a^3)}\right]$$

Then insert into the formula for a z value

Example 5

So, if an article presented a p value of .021 (N=60), then the calculation would be as follows

$$a = \sqrt{\ln(1/.021^2)}$$

$$a = \sqrt{2267.57}$$

$$a = 7.7265$$

$$z = 7.7265\left[\frac{(2.515517)(.802853 \times 7.7265)(.010328 \times 59.698)}{1 - (1.432788 \times 7.7265)(.189269 \times 59.698)(.001308 \times 461.262)}\right]$$

z = 7.7265 (9.621/-74.41)

z =.99

Then, calculating an r from a z value

$$r = \sqrt{z^2/N}$$

$$r = \sqrt{.99^2 / 60}$$

r = .127

Using the formula for an effect size from r

$$d = \sqrt{\frac{4(.127)^2}{1-(.127)^2}}$$

d = .256

Calculating an effect size for a cross-over study

Equation 9.

$$d = \frac{\overline{X}_e - \overline{X}_c}{SDc\sqrt{1-r_{ec}{}^2}}$$ where r_{ec} is the correlation between the experimental and

control. For cross-over studies the correlation is between pre-post because experimental and control subjects are the same.

Example 6.

If the lowest correlation between pre-post measures is reported to be or calculated to be .45, experimental mean of 3.69 and a control mean of 2.54, A pooled standard deviation of 3.65. (Same numbers from Example 1.) Results are as follows,

$$d = \frac{3.69 - 2.54}{3.65\sqrt{1-.45^2}}$$

d = 1.15/3.65 (.89)

$d = 1.15/3.25$

$d = .35$ (to note difference adding correlation makes compare to effect size calculation from Example 1. where $d = .308$)

Calculating pooled standard deviations

Equation 10

$$Sp = \sqrt{\frac{(n1 - 1)s1^2 + (n2 - 1)s2^2}{n1 + n2 - 2}}$$

Example 7

So, if a article presents with a SD for a treatment group of 3.44 (N = 80) and a SD for a control group of 3.85 (N = 80), the calculation would be

$$Sp = \sqrt{\frac{79(3.44)^2 + 79(3.85)^2}{79 + 79 - 2}}$$

Sp = 3.65

Calculating an Odds Ratio

Equation 11

$OR = \frac{ad}{bc}$ where a = positive outcomes for experimental group , b = positive

outcomes for control, c = negative outcomes for the experimental group, and d = negative outcomes from the control groups. a+b+c + d = N.

Example 8.

If an article presents the information that out of 30 experimental subjects 22 achieved at least a 50% reduction in symptoms and 10 out of 30 in the control group achieved the same criteria, the calculation would be as follows.

a = 22 , b = 10, c = 8, d = 20

OR =(22)(20)/(10)(8)

OR = 5.5

Calculating Cochrane's Q

Equation 12.

$Q_{\hat{w}} = \sum \hat{w}(Y_i - \hat{\mu})^2$ where w= the inverse of the variance (or weight), Yi = the effect size of each trial, and μ = the overall effect estimated from the meta-analysis.

Example 9.

If a meta-analysis is being performed on three studies and effect sizes have been calculated on continuous data the following information would be available.

Study 1 d = .40 variance = .11 weight (inverse of the variance) = 9.5

Study 2 d = .51 variance = .10 weight (inverse of the variance) = 9.3

Study 3 d = .51 variance = .17 weight (inverse of the variance) = 5.8

If these numbers used to calculate an overall effect size with the REM model (using Meta-Analysis ® computer program) the population effect size estimate is found to be .47. Using this as the population estimate of effect a Cochrane's Q can be calculated.

$$Q_w = 9.5(.40-.47)^2 + 9.3(.51-.47)^2 + 5.8(.51-.47)^2$$

$$Q_{\hat{w}} = 9.5(.0049) + 9.3(.0016)+5.8(.0016)$$

$$Q_{\hat{w}} = .069$$

A very small Cochrane's Q which indicates very little variation between studies.

Appendix G: Regression Analyses and Correlation Results for Propranolol

Table G1

Summary of Regression Analysis for Variables Predicting Treatment Outcome Results for Propranolol

Variable	B	SE B	*p*
Entered			
Study type	.483	.172	.011
Excluded	**B**	***t***	***p***
Comparison group	.28	.138	.892
Control group type	.188	.999	.330
Length of Tx	-.002	-.008	.994
Outcome measure	-.263	-1.453	.162
Quality score (Jadad)	-.224	-1.196	.246
Quality rating	-.266	-1.431	.168
Mortality rate	-.072	-.378	.710
Percent female	.036	.189	.852
Publication year	.188	.883	.388

Note. $R^2 = .273$, $F = 7.89$.

Table G2

Correlation Between Variables and Unbiased Effect Sizes for Propranolol

Variable	Correlation (*r*) with unbiased effect size	Variable	Correlation (*r*) with unbiased effect size
Comparison group	-.42	Quality rating	-.367
Control group type	.080	Mortality rate	-.058
Length of Tx	-.272	Percent female	-.045
Outcome measure	-.257	Publication year	.109
Quality score(Jadad)	-.308	Study type	.514**

Note. ** Correlation is significant at that 0.01 level (2-tailed).

Appendix H: Regression Analyses and Correlation Results for Flunarizine

Table H1

Summary of Regression Analysis for Variables Predicting Treatment Outcome Results for Flunarizine

Variable	B	SE B	*p*
Entered			
Treatment length	.772	.014	.001
Excluded	**B**	***t***	***p***
Comparison group	-.231	-.232	.243
Control group type	.008	-.037	.971
Outcome measure	-.033	-.167	.871
Quality score (Jadad)	.191	1.019	.330
Quality rating	.320	1.887	.086
Mortality rate	-.130	-.624	.545
Percent female	-299	-1.699	.117
Publication year	-.533	-3.037	.011
Study type	.220	1.165	.269

Note. R^2=.587, F = 17.074.

Table H2

Correlation Between Variables and Unbiased Effect Sizes for Flunarizine

Variable	Correlation (*r*) with unbiased effect size	Variable	Correlation (*r*) with unbiased effect siz
Comparison group	-.403	Quality rating	-.211
Control group type	.303	Mortality rate	-.417
Length of Tx.	.733**	Percent female	-.135
Outcome measure	-.182	Publication year	.082
Quality score(Jadad)	.061	Study type	.399

Note. ** Correlation is significant at that 0.01 level (2-tailed).

Appendix I: Regression Analyses and Correlation

Results for Divalproex Sodium

None of the variables met the inclusionary criteria for a stepwise regression ($F < .05$) thus variables were forced in through the enter method to obtain values.

Table I1

Summary of Regression Analysis for Variables Predicting Treatment Outcome Results for Divalproex Sodium

Variable	B	SE B	*p*
Entered			
None			
Excluded	**B**	***t***	***p***
Comparison group	N/A	N/A	N/A
Control group	N/A	N/A	N/A
Quality score (Jadad)	-.234	-.627	.564
Quality rating	-.332	-.863	.437
Outcome measure	N/A	N/A	N/A
Mortality rate	-.356	-1.150	.314
Percent female	-.452	-1.376	.241
Publication year	-.037	-.106	.921
Treatment length	4.071	.646	.553
Study type	.324	1.013	.368

Note. R^2=1.00. Comparison group, control group, and outcome measure deleted from regression because they were constants or nearly constants could not be computed.

Table I2

Correlation Between Variables and Unbiased Effect Sizes for Divalproex Sodium

Variable	Correlation (*r*) with unbiased effect size	Variable	Correlation (*r*) with unbiased effect size
Comparison group	.72	Quality rating	-.631
Control group type	-.72	Mortality rate	-.168
Length of Tx	-.223	Percent female	-.698
Outcome measure	N/A	Publication year	.165
Quality score (Jadad)	-.515	Study type	.117

Note. ** Correlation is significant at that 0.01 level (2-tailed). Outcome measure deleted because it was a constant.

Appendix J:Regression Analyses and Correlation

Results for Thermal Biofeedback

None of the variables met the inclusionary criteria for a stepwise regression ($F < .05$) thus variables were forced in through the enter method to obtain values.

Table J1

Summary of Regression Analysis for Variables Predicting Treatment Outcome Results for Thermal Biofeedback

Variable	B	SE B	*p*
Entered			
None			
Excluded	**B**	***t***	***p***
Comparison group	-.176	-3.920	.030
Control group type	2.641	.705	.532
Outcome measure	-.233	-2.210	.114
Quality score (Jadad)	-.686	-3.012	.057
Quality rating	-6.23	-.310	.777
Mortality rate	3.446	-2.357	.100
Percent female	8.011	1.650	.198
Publication year	3.355	.344	.754
Treatment length	8.341	.035	.974
Study type	N/A	N/A	N/A

Note. $R^2 = .928$, $F = 3.876$. Variable study type removed because is was a constant.

Table J2

Correlation Between Variables and Unbiased Effect Sizes for Thermal Biofeedback

Variable	Correlation (*r*) with unbiased effect size	Variable	Correlation (*r*) with unbiased effect size
Comparison group	-.072	Quality rating	-.249
Control group type	.495	Mortality rate	-.222
Length of Tx.	.159	Percent female	-.440
Outcome measure	-.411	Publication year	.221
Quality score (Jadad)	-.241	Study type	N/A

Note. Study type removed because it was a constant.

Appendix K: Correlation Tables for Relaxation Therapy

Table K1

Correlations Between Variables and Unbiased Effect Sizes for Relaxation Therapy

Variable	Correlation (*r*) with unbiased effect size	Variable	Correlation (*r*) with unbiased effect size
Comparison group	-.603	Quality rating	-.698
Control group type	N/A	Mortality rate	.159
Length of Tx.	.459	Percent female	.057
Outcome Measure	.603	Publication year	-.389
Quality score (Jadad)	-.698	Study type	N/A

Note. Control group and study type deleted because they are constants.

Appendix L: Post-Hoc Multiple Comparisons with Bonferroni Corrections

Table L1

Post-Hoc Comparisons of Treatments for Continuous Effect Sizes

Treatments compared	*t*	Standard error of the difference	*p* value
Propranolol vs. Flunarizine	0.000	0.023	1.000
Propranolol vs. Thermal Bio.	3.113	0.026	0.008
Propranolol vs. Relaxation	1.452	0.048	0.157
Propranolol vs. Thermal + Cog.	0.760	0.053	0.453
Flunarizine vs. Thermal Bio.	2.522	0.032	0.017
Flunairizne vs. Relaxation	1.175	0.060	0.256
Flunarizine vs. Thermal + Cog.	0.619	0.065	0.544
Thermal vs. Relaxation	2.238	0.067	0.038
Thermal vs. Thermal + Cog.	1.6215	0.074	0.123
Relax vs. Thermal + Cog.	0.196	0.153	0.852

Note. Bonferroni Correction criteria is $p < .005$. ANOVA $F = 2.9024$, $p = 0$ .0297.

Table L2

Post-Hoc Bonferroni Comparisons of Treatments for Continuous Effect Sizes

Treatments compared	*t*	Standard error of the difference	*p* value
Propranolol vs. Flunarizine	0.060	0.148	0.562
Propranolol vs. Divalproex Sodium	0.745	0.094	0.475
Propranolol vs. Thermal Bio.	0.000	0.019	1.000
Propranolol vs. EMG Bio.	3.414	0.261	0.014
Propranolol vs.EMG + Thermal Bio.	2.384	0.191	0.048
Flunarizine vs. Divalproex Sodium	0.745	0.215	0.483
Flunarizine vs. Thermal Bio.	0.603	0.149	0.565
Flunarizine vs. EMG Bio.	1.946	0.503	0.1468
Flunarizine vs. EMG + Thermal Bio.	1.012	0.379	0.368
Divalproex Sodium vs. Thermal Bio.	0.210	0.095	0.837
Divalproex Sodium vs. EMG Bio.	2.682	0.339	0.043
Divalproex Sodium vs. EMG + Thermal Bio.	1.737	0.259	0.132
Thermal Bio. vs. EMG Bio.	3.402	0.262	0.014
Thermal Bio. vs EMG +Thermal Bio.	2.375	0.198	0.049
EMG Bio. vs. EMG + Thermal Bio.S	2.407	0.565	0.095

Note. Bonferroni correction criteria is $p < .003$. ANOVA $F = 3.177$, $p = 0.314$.

VITA

Kristi Lowe Stewart

School Address:
Department of Psychology
2810 Old Main Hill
Logan, Utah 84322-2810

1492 East Spring Lane #317
Salt Lake City, Utah, 84117
(801) 277-5271
(801) 412-2563

Education

PhD Internship	Salt Lake City Veterans Administration Medical Center Medical Psychology/Substance Abuse/ Outpatient Mental Health **Full APA Accreditation**
MS	Utah State University, Logan, Utah, June, 1998. Clinical/Counseling Psychology *School Psychology Certified* **Full APA Accreditation**
BS	University of Utah, June 1994 Major: Psychology

Clinical Experience and Employment

Adult Psychology

2001-present	***Therapist,*** Jordan Family Education Center, Midvale, Utah. Family and couples intervention.
2001-present	***Instructor,*** Jordan Family Education Center. Midvale, Utah. *Calming the Storm*, Anger management for adults.
2001-present	***Psychology Resident***, University of Utah Neuropsychiatric Institute. PRN intervention Child/Adolescent/Family/Adult.
1999-2000	***Therapist,*** Veterans Administration Hospital, Salt Lake City, Utah. Individual/Couples therapy, Inpatient/Outpatient, and Substance Abuse. Focus on individuals who are substance abuses with personality disorders and/or trauma histories
1999	***Therapist,*** Bear River Mental Health, Logan, Utah. Individual and child assessment and treatment. General mental health, low income and medicaid/medicare clients.
1996 - 1999	***Therapist,*** Psychology Community Clinic, Utah State University. Individual, couple, family and child therapy. Focus on general mental health and personality disorders.

1997 - 1999 ***Therapist,*** Student Counseling Center, Utah State University.
Individual assessment and treatment of students. General mental health and crisis work.

Behavioral Medicine/Substance Abuse

2001-present ***Psychology Resident***, University of Utah Neuropsychiatric Institute.
PRN intervention Child/Adolescent/Family/Adult.

2000-2001 ***Psychology Resident***, University of Utah Neuropsychiatric Institute.
Child/Adolescent/Family/Adult inpatient hospital: neuorpsychological assessment, psychological assessment, intervention, detox, pain management, crisis intervention.

1999 – 2000 ***Psychology Intern,*** Salt Lake City Veterans Administration Medical Center.
Medical Psychology Rotation: Neuorpsychological assessment, cardiac transplant and presurgical evaluations, pain management, adjustment to illness, competency evaluations, rehabilitation, interdisciplinary team member, consultation.

Inpatient/Outpatient Mental Health and Substance Abuse: Drug and alcohol assessment/treatment/detox, inpatient assessment/treatment, PTSD, geriatric psychology, adult mental health.

Child Psychology

2001-present ***School Psychologist***, Jordan School District, Sandy, Utah.
Assessment and treatment for children under the age of 15. Family/Individual/Group.

2001-present ***Psychology Resident***, University of Utah Neuropsychiatric Institute.
PRN intervention Child/Adolescent/Family/Adult.

2001-present ***Therapist,*** Jordan Family Education Center, Midvale, Utah.
Family and couples intervention.

2000-2001 ***Psychology Resident***, University of Utah Neuropsychiatric Institute.
Child/Adolescent/Family/Adult inpatient hospital: neuorpsychological assessment, psychological assessment, intervention, crisis intervention..

1997 - 1998 ***Child Therapist,*** Clinical Services, Center for Persons with Disabilities, Utah State University.
Psychological assessment, diagnosis and treatment planning for children and adolescents with psychological and behavioral disorders. Group and individual

1996 -1998 ***Child/Adolescent Therapist***, Child Evaluation and Treatment Center, Logan, Utah.
Treatment, evaluation, and report writing for psychological disorders in children, adolescents, and families.

1996 - 1997 ***School Psychology Intern,*** Davis County School District. Psychological and psychoeducational evaluation of elementary and secondary students for classification and placement purposes.

1996 - 1997 ***Child Psychological Assessment***, Community Family Partnership, Center for Persons with Disabilities, Utah State University. Psychological assessment of children age one month to nine years with focus on developmental and intellectual functioning.

1991-1993 ***Residential Therapist***, The Children's Center, Kearns, Utah.
Behavioral management, social, emotional, intellectual and physical development interventions for children under the age of seven.

Administrative Experience

2003-present ***Supervisor,*** Jordan School District, Sandy, Utah.
Individual supervision of medical students specializing in child psychiatry.

2001-present ***Supervisor***, Jordan School District, Sandy, Utah.
Individual supervision of University of Utah practicum students in school psychology.

1999-2000 ***Supervisor,*** Veterans Administration Hospital, Salt Lake City, Utah.
Individual and group supervision of University of Utah practicum students in the Substance Abuse Counseling Program

1998-1999 ***Supervisor,*** Counseling Center, Utah State University.
Individual and group supervision of a group of undergraduate students in the Peer Counselor Program.

1997-1998 ***Graduate Student Representative,*** Utah State University.
Program development, facilitation of student/faculty issues, and graduate student selection committee.

1996-1997 ***Graduate Student Representative Faculty Selection Committee,*** Utah State University.
Organization of faculty candidate meetings, facilitation of student interviews and mock clinical interviews.

1995-1996 ***Assistant to the Clinic Director***, Psychology Community Clinic, Utah State University.
Supervision of clinical testing, data base management, community outreach, and marketing.

Research Experience

Current ***Doctoral Dissertation,*** Utah State University, *Pharmacological and Behavioral Treatments for Migraine Headaches: A Meta Analysis*. Compares the effectiveness of available prophylactic treatments for migraine headaches. Supervisor: Kevin S. Masters, PhD

1997-1998 ***Masters Thesis,*** Utah State University, Stewart, K.L. *Parenting Styles and Internalizing Behavioral Problems in Children : An Exploratory Study. Assessment of the associations between parenting behaviors and childhood internalizing problems.* Thesis Chair: Gretchen A. Gimpel Ph.D.

1997 ***Research Assistant,*** Utah State University. A state wide survey of teachers perceptions of children with behavioral disorders and learning disabilities. Supervisor: Gretchen A. Gimpel Ph.D.

1991-1994 ***Research Assistant,*** University of Utah. Volunteer bias in human sexuality research, The effects of erotic films on stereotypical attitudes, MMPI-2 validity studies, and Developmental experiences' impact on adult problems solving abilities. Supervisor: Donald Strassberg Ph.D.

Publications

Strassberg, D & Lowe, K (1995). Volunteer Bias in Human Sexuality Research. *The Archives of Sexual Behavior, 24*, 369-382.

Presentations

Lowe Stewart, K.L. (2002). Students with disabilities- accommodating 504 students. Lecture at the Utah Counseling Association Conference, Ogden, Utah.

Stewart, K.L., & Shearer, D.S. (1999). *The Impact of Parenting Style on Elementary School Student Behavior.* Paper Presented at the 1999 annual National Association of School Psychologists, Las Vegas Nevada.

Shearer, D.S., & Stewart, K.L. (1998). *The relationship between Attention Deficit/Hyperactivity Disorder, the TOVA and the ADDES: Implications for Diagnosis.* Presented at the 1988 annual Virginia Beach Conference, Virginia Beach, Virginia.

Stewart, K.L, & Shearer, D.S. (1998). *Internalizing symptoms in school-aged children: Stability, recognition, and measurement.* Presented at the 1998 annual Utah Association of School Psychologists Conference, Salt Lake City, Utah.

Stewart, K.L., & Gimpel, G.A. (1997). *Teachers perceptions of Attention Deficit/Hyperactivity Disorder.* Presented at the 1997 annual National Association of School Psychologists Conference, Anehiem, California.

Stewart, K.L. & Strassberg, D.S. (1994). *The use of student populations in human sexuality research.* Presented at the annual University of Utah Student Research Conference, Salt Lake City, Utah.

Organizational Activities

American Psychological Association
National Association of School Psychologists
Utah Psychological Association

Honors

Psychology Department Fellowship-Utah State University 1997
Dean's Honor List, University of Utah- 1993-1994
Psychology Department Scholarship, University of Utah, 1993-1994

References

Scott Hill, PhD, Staff Psychologist, Salt Lake City Veterans Administration Medical Center, Intern Supervisor
500 Foothill Blvd., Salt Lake City, Utah 84148
(801) 582-1565 x2284

Kelly Lundberg, PhD, Director of Substance Abuse Services, University of Utah.

Kay Koellner, PhD, Staff Psychologist, Salt Lake City Veterans Administration Medical Center, Intern Supervisor
500 Foothill Blvd., Salt Lake City, Utah 84148
(801) 582-1565 x2284

Kevin Masters, PhD, Associate Professor, Utah State University, Dissertation Chair
2810 Old Main Hill, Department of Psychology, Logan, Utah 84322-2810
(435) 797-1463